



# **NAVAL POSTGRADUATE SCHOOL**

**MONTEREY, CALIFORNIA**

## **THESIS**

**STATISTICAL ANALYSIS OF ENSEMBLE FORECASTS  
OF TROPICAL CYCLONE TRACKS OVER THE  
NORTHWEST PACIFIC OCEAN**

by

David R. Marino

September 2012

Thesis Advisor:  
Second Reader:

Patrick A. Harr  
Joshua P. Hacker

**Approved for public release; distribution unlimited**

THIS PAGE INTENTIONALLY LEFT BLANK

<b>REPORT DOCUMENTATION PAGE</b>			<i>Form Approved OMB No. 0704-0188</i>	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington DC 20503.				
<b>1. AGENCY USE ONLY (Leave blank)</b>		<b>2. REPORT DATE</b> September 2012	<b>3. REPORT TYPE AND DATES COVERED</b> Master's Thesis	
<b>4. TITLE AND SUBTITLE</b> Statistical Analysis of Ensemble Forecasts of Tropical Cyclone Tracks over the Northwest Pacific Ocean			<b>5. FUNDING NUMBERS</b>	
<b>6. AUTHOR(S)</b> David R. Marino				
<b>7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)</b> Naval Postgraduate School Monterey, CA 93943-5000			<b>8. PERFORMING ORGANIZATION REPORT NUMBER</b>	
<b>9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES)</b> N/A			<b>10. SPONSORING/MONITORING AGENCY REPORT NUMBER</b>	
<b>11. SUPPLEMENTARY NOTES</b> The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. IRB Protocol number _____N/A_____.				
<b>12a. DISTRIBUTION / AVAILABILITY STATEMENT</b> Approved for public release; distribution is unlimited			<b>12b. DISTRIBUTION CODE</b> A	
<b>13. ABSTRACT (maximum 200 words)</b>  The skill of three ensemble prediction systems (EPS) is evaluated to focus on tropical cyclone (TC) track forecasting over the North Pacific. Probability ellipses are defined to represent ensemble spread and encompass 68 % of the ensemble members. The ellipses are centered on the ensemble mean forecast position. Forecast reliability is defined as whether the verifying position is within the ellipse 68% of the time. A statistical analysis of uncertainty in TC track forecasts examines the attributes of reliability and resolution of each EPS. The European Center for Medium-Range Forecasts (ECMWF) EPS had the highest degree of reliability and resolution. The sizes and shapes of the EPS ellipses varied with TC track characteristics. This suggests that EPS-based probability ellipses may provide value in identifying uncertainty with respect to likely TC track forecast errors.				
<b>14. SUBJECT TERMS</b> Tropical Cyclone Track Errors, Ensemble Prediction Systems, Joint Typhoon Warning Center forecasts			<b>15. NUMBER OF PAGES</b> 83	
			<b>16. PRICE CODE</b>	
<b>17. SECURITY CLASSIFICATION OF REPORT</b> Unclassified	<b>18. SECURITY CLASSIFICATION OF THIS PAGE</b> Unclassified	<b>19. SECURITY CLASSIFICATION OF ABSTRACT</b> Unclassified	<b>20. LIMITATION OF ABSTRACT</b> UU	

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)  
Prescribed by ANSI Std. Z39-18

THIS PAGE INTENTIONALLY LEFT BLANK

**Approved for public release; distribution unlimited**

**STATISTICAL ANALYSIS OF ENSEMBLE FORECASTS OF TROPICAL  
CYCLONE TRACKS OVER THE NORTHWEST PACIFIC OCEAN**

David R. Marino  
Lieutenant Commander, United States Navy  
B.S. Oceanography, United States Naval Academy, 2002

Submitted in partial fulfillment of the  
requirements for the degree of

**MASTER OF SCIENCE IN METEOROLOGY AND OCEANOGRAPHY**

from the

**NAVAL POSTGRADUATE SCHOOL  
SEPTEMBER 2012**

Author: David R. Marino

Approved by: Patrick A. Harr  
Thesis Advisor

Joshua P. Hacker  
Second Reader

Wendell A. Nuss  
Chair, Department of Meteorology

THIS PAGE INTENTIONALLY LEFT BLANK

## **ABSTRACT**

The skill of three ensemble prediction systems (EPS) is evaluated to focus on tropical cyclone (TC) track forecasting over the North Pacific. Probability ellipses are defined to represent ensemble spread and encompass 68 % of the ensemble members. The ellipses are centered on the ensemble mean forecast position. Forecast reliability is defined as whether the verifying position is within the ellipse 68% of the time. A statistical analysis of uncertainty in TC track forecasts examines the attributes of reliability and resolution of each EPS. The European Center for Medium-Range Forecasts (ECMWF) EPS had the highest degree of reliability and resolution. The sizes and shapes of the EPS ellipses varied with TC track characteristics. This suggests that EPS-based probability ellipses may provide value in identifying uncertainty with respect to likely TC track forecast error.

THIS PAGE INTENTIONALLY LEFT BLANK



# TABLE OF CONTENTS

<b>I.</b>	<b>INTRODUCTION.....</b>	<b>1</b>
<b>A.</b>	<b>MOTIVATION .....</b>	<b>1</b>
<b>B.</b>	<b>OBJECTIVE .....</b>	<b>5</b>
<b>II.</b>	<b>BACKGROUND .....</b>	<b>7</b>
<b>A.</b>	<b>JOINT TYPHOON WARNING CENTER OPERATIONAL METHODS TO DEFINE TROPICAL CYCLONE TRACK FORECAST UNCERTAINTY .....</b>	<b>7</b>
1.	Official Forecast Track with Area of Uncertainty .....	7
2.	Wind Speed Probability Swath.....	8
3.	The Goerss Predicted Consensus Error (GPCE).....	9
4.	Goerss Predicted Consensus Error Along- and Across-track (GPCE-X).....	11
<b>B.</b>	<b>ENSEMBLE PREDICTION SYSTEMS (EPS) .....</b>	<b>11</b>
1.	European Center for Medium Range Weather Forecasts (ECMWF) .....	12
2.	United Kingdom Meteorological Office (UKMO) .....	13
3.	Japan Meteorological Agency Global Spectral Model (JMA/GSM) .....	14
<b>III.</b>	<b>METHODOLOGY .....</b>	<b>15</b>
<b>A.</b>	<b>DATA .....</b>	<b>15</b>
1.	Data Source.....	15
2.	Data Format .....	18
3.	Data Homogeneity.....	19
4.	Region and Sub-regions.....	19
5.	Developing the EPS Ellipse .....	21
<b>B.</b>	<b>STATISTICAL ANALYSES .....</b>	<b>23</b>
1.	Probability within Spread .....	23
2.	Ellipse Reliability .....	24
3.	Mean Area Difference (MAD) .....	24
<b>IV.</b>	<b>ANALYSIS AND RESULTS .....</b>	<b>25</b>
<b>A.</b>	<b>OVERVIEW .....</b>	<b>25</b>
<b>B.</b>	<b>ENSEMBLE-MEAN TRACK ERRORS.....</b>	<b>25</b>
1.	FTE.....	25
2.	ATE .....	30
3.	XTE .....	35
<b>C.</b>	<b>PROBABILITY WITHIN SPREAD.....</b>	<b>40</b>
<b>D.</b>	<b>ELLIPSE RELIABILITY .....</b>	<b>42</b>
<b>E.</b>	<b>MEAN AREA DIFFERENCE (MAD).....</b>	<b>44</b>
<b>F.</b>	<b>SUMMARY .....</b>	<b>45</b>
<b>G.</b>	<b>CASE STUDY TYPHOONS NANMADOL AND SONGDA .....</b>	<b>47</b>
1.	Typhoon Songda.....	49

a.	<i>Typhoon Songda 1200 UTC 23 May 2011 .....</i>	<i>49</i>
b.	<i>Typhoon Songda 1200 UTC 24 May 2011 .....</i>	<i>51</i>
c.	<i>Typhoon Songda 1200 UTC 25 May 2011 .....</i>	<i>52</i>
2.	<b>Typhoon Nanmadol .....</b>	<b>52</b>
a.	<i>Typhoon Nanmadol 1200 UTC 23 August 2011.....</i>	<i>53</i>
b.	<i>Typhoon Nanmadol 1200 UTC 24 August 2011.....</i>	<i>54</i>
c.	<i>Typhoon Nanmadol 1200 UTC 25 August 2011.....</i>	<i>55</i>
3.	<b>Summary.....</b>	<b>57</b>
V.	<b>CONCLUSIONS AND RECOMMENDATIONS.....</b>	<b>59</b>
A.	<b>CONCLUSIONS .....</b>	<b>59</b>
B.	<b>RECOMMENDATIONS.....</b>	<b>59</b>
	<b>LIST OF REFERENCES.....</b>	<b>61</b>
	<b>INITIAL DISTRIBUTION LIST .....</b>	<b>63</b>

## LIST OF FIGURES

Figure 1.	The JTWC track forecasts (squares) and best-track (tropical cyclone symbol) for STY 14W (From: JTWC 2012a).....	4
Figure 2.	The JTWC Warning Graphic Legend. The solid black TC symbol defines the location of a TC, with winds greater than 64 knots. The last solid black TC symbol is the current position, and the pink cyclone symbols are forecasted positions. The pink circle, around the pink cyclone symbol, is the 64 knot wind radii. The shaded area of uncertainty is the 34 knot wind radii plus the average forecast track error (From: JTWC 2012b).....	8
Figure 3.	Wind speed probability graphic that depicts the likelihood of 50-kt winds will occur during the next 120 h issued 24 August 2011. Legend at the top of the box explains color scale representing the probabilities of 50-kt winds. The last black tropical cyclone symbol represents the position of the storm when the graphic was created (From: JTWC 2012c).....	9
Figure 4.	Predicted 70% confidence radius (solid circle) of the 120-h CONU forecast for Hurricane Isabel on 0000 UTC 13 September 2003. The individual model tracks used to create the CONU consensus model are shown. Notice the GPCE circle is much smaller than the 120-h radius (dotted circle) used by the NHC potential 5-day track area graphic (From: Goerss 2007) .....	10
Figure 5.	Schematic definition of forecast-track error (FTE), cross-track error (XTE), and along-track error (ATE) (From: JTWC 2012a) .....	11
Figure 6.	Tropical cyclones that occurred during (a) April through early September 2009, and (b) September through December 2009 (From: JTWC 2012a) .....	16
Figure 7.	Tropical cyclones that occurred during (a) April through early September 2010, and (b) September through December 2010 (From: JTWC 2012a) .....	17
Figure 8.	Tropical cyclones that occurred during (a) April through early September 2011, and (b) September through December 2011 (From: JTWC 2012a) .....	18
Figure 9.	Geographic sub-regions of the WPAC region used to group Ensemble Prediction System (EPS) forecast track data .....	20
Figure 10.	The probability ellipse (red) contains 68% of the ensemble members (red dots) and is centered on the ellipse mean position. $y_1$ and $y_2$ are the coordinate system by which the ensemble members are rotated around. The orientation of the ellipse is defined by $\phi$ .....	22
Figure 11.	Average FTE for each EPS over the entire WPAC from 2009–2011. Plus and minus one standard deviation about the FTE is represented by a vertically oriented line at each 12-h forecast from 0–120 h .....	26
Figure 12.	As in Figure 11, except for the MTR region.....	27
Figure 13.	As in Figure 11, except for the RCR region .....	28
Figure 14.	As in Figure 11, except for the SCSR region.....	29
Figure 15.	As in Figure 11, except for the MLR region.....	30

Figure 16.	Average ATE for each of the EPS over the entire WPAC from 2009–2011. A plus and minus one standard deviation about the mean ATE is represented by the vertical line at each 12-h forecast from 0–120 h .....	31
Figure 17.	As in Figure 16, except for the MTR region.....	32
Figure 18.	As in Figure 16, except for the RCR region .....	33
Figure 19.	As in Figure 16, except for the SCSR region.....	34
Figure 20.	As in Figure 16, except for the MLR region.....	35
Figure 21.	Average XTE for each of the EPS of the entire WPAC from 2009–2011. A plus and minus one standard deviation about the mean XTE is represented by a vertical line at each 12 h forecast from 0–120 h .....	36
Figure 22.	As in Figure 21, except for the MTR region.....	37
Figure 23.	As in Figure 21, except for the RCR region .....	38
Figure 24.	As in Figure 21, except for the SCSR region.....	39
Figure 25.	As in Figure 21, except for the MLR region.....	40
Figure 26.	The PWS for each EPS over the entire WPAC for the 2009–2011 seasons....	41
Figure 27.	The PWS for each EPS over each sub-region of the WPAC for the 2009–2011 seasons .....	42
Figure 28.	The ellipse reliability for each EPS for the entire WPAC .....	43
Figure 29.	The ellipse reliability for each EPS over the each sub-region of the WPAC for the 2009–2011 seasons as defined in Figure 28.....	44
Figure 30.	The Mean Area Difference (MAD) of UKMO or JMA EPS and compared to ECMWF across the WPAC. Positive values indicate the EPS ellipses are smaller than the ECMWF circle for each forecast interval. The values above the forecast interval are the number of EPS forecasts included .....	45
Figure 31.	All JTWC track forecasts for Typhoon Nanmadol. The red dashed lines correspond to each JTWC forecast track. The best-track positions are in black (From: JTWC 2012a) .....	48
Figure 32.	All JTWC track forecasts for Typhoon Songda. The red dashed lines correspond to each JTWC forecast track. The best-track positions are in black (From: JTWC 2012a) .....	49
Figure 33.	The TC forecast track ellipses of each EPS for Typhoon Songda 1200 UTC 23 May 2011. Each ellipse signifies a 12 h forecast interval and is colored to match the individual EPS as defined in the legend at the top right. The large dot inside each ellipse is the corresponding ensemble-mean forecast position. The best-track positions are in black .....	50
Figure 34.	The TC forecast track ellipses of each EPS for Typhoon Songda 1200 UTC 24 May 2011, as defined in Figure 33 .....	51
Figure 35.	The TC forecast track ellipses of each EPS for Typhoon Songda 1200 UTC 25 May 2011, as defined in Figure 33 .....	52
Figure 36.	The TC forecast track ellipses of each EPS for Typhoon Nanmadol 1200 UTC 23 August 2011. Each ellipse signifies a 12 h forecast interval and is colored to match the individual EPS as defined in the legend at the top right. The large dot inside each ellipse is the corresponding ensemble-mean forecast position. The best-track positions are in black .....	54

Figure 37.	The TC forecast track ellipses of each EPS for Typhoon Nanmadol 1200 UTC 24 August 2011, as defined in Figure 36 .....	55
Figure 38.	The TC forecast track ellipses of each EPS for Typhoon Nanmadol 1200 UTC 25 August 2011, as defined in Figure 36 .....	56

THIS PAGE INTENTIONALLY LEFT BLANK

## LIST OF TABLES

Table 1.	Total number of Tropical Cyclones included in this thesis for the WPAC and each sub-region .....	20
----------	--	----

THIS PAGE INTENTIONALLY LEFT BLANK



## **LIST OF ACRONYMS AND ABBREVIATIONS**

AGBOM	Australia Government Bureau of Meteorology
AOR	Area of Responsibility
ATCF	Automated Tropical Cyclone Forecast
ATE	Along-Track Error
AVNI	Global Forecast System Aviation Interpolated
CONU	Consensus Version U
COR	Condition of Readiness
CXML	Cyclone XML
DoD	Department of Defense
ECMWF	European Center for Medium-range Weather Forecasts
EDA	Ensemble of Data Assimilations
EEM	ECMWF Ensemble-mean
EPS	Ensemble Prediction System
ETKF	Ensemble Transform Kalman Filter
FTE	Forecast-Track Error
GE	Grand Ensemble
GFDI	GFDL Interpolated
GFDL	Geophysical Fluid Dynamics Laboratory
GFNI	Navy GFDL Interpolated
GFS	Global Forecast System
GPCE	Goerss Predicted Consensus Error
GPCE-X	Goerss Predicted Consensus Error-Along/Across-track
GSM	Global Spectral Model
h	hour
JMA	Japan Meteorological Agency
JTWC	Joint Typhoon Warning Center
kt	knot(s)
MAD	Mean Area Difference
MCP	Monte Carlo Probability
MLR	Mid-Latitude Region

MOGREPS	UKMET Office Global and Regional Ensemble Prediction System
MTR	Monsoon Trough Region
NCEP	National Centers for Environmental Prediction
NGPI	NOGAPS Interpolated
NHC	National Hurricane Center
NOGAPS	Navy's Operational Global Atmospheric Prediction System
NOOC	Naval Oceanography Operational Command
PWS	Probability within Spread
SV	Singular Vector
TC	Tropical Cyclone
TIGGE	THORPEX Interactive Grand Global Ensemble
TUTT	Tropical Upper Tropospheric Trough
UKMET	United Kingdom Meteorological Office
UKMI	UKMET Global Model Interpolated
UKMO	United Kingdom Meteorological Office
RCR	Re-curve Region
SCRS	South China Sea Region
UTC	Coordinated Universal Time
WPAC	Western Pacific
XTE	Cross-Track Error

## **ACKNOWLEDGMENTS**

Thank you to my amazing wife, Giovanna who always supports me. A special emphasis on this year where she supported me through helping me heal a broken neck, the thesis process, and completion of my master's degree.

A special acknowledgment to my advisor, Professor Pat Harr, for all of his efforts in guiding me through the thesis process, providing a wealth of knowledge, and keeping me on track. Thank you to Professor Joshua Hacker for his thorough editorial support and insight. A word of gratitude must be offered to LCDR Doug Pearman for the motivation, clear guidance, needed criticisms, and insight, which kept me focused on completion of this thesis. I would also like to thank my officemates, Louis and Rob, for keeping the atmosphere in our office light and amusing through this process.

THIS PAGE INTENTIONALLY LEFT BLANK

# **I. INTRODUCTION**

## **A. MOTIVATION**

Tropical Cyclones (TCs) are among the most damaging and costly synoptic-scale weather phenomena that threaten military aviation and maritime assets and DoD installations along coastlines and on island chains. In the North Pacific basin, there are 35 DoD installations, of which 25 are located in the Western Pacific (WPAC) region. Also, the WPAC experiences the highest number of TCs of all ocean basins. Since these installations and their assigned Areas of Responsibility (AOR) overlap with the region of known maximum climatological TC occurrence, there are major fiscal and safety concerns to be addressed by installation and unit commanders when the threat of a TC exists.

During the period from 2009–2011, there were 54 named storms in the WPAC. Each of these storms had the potential to impact operations at sea, in the air, and on land. Monetary impacts of these storms can reach significant levels due to the costs of moving ships and aircraft, evacuating personnel, and preparing installations for damage prevention. Potential losses of property and life caused by a TC could be much more significant without proper preparation. It is the responsibility of each installation, or unit commander to make a highly informed decision to mitigate the risk of sustaining damages.

The regional Naval Fleet Oceanographer, with the support of his staff and AOR forecasting center, use numerical models, official AOR forecasts, and TC products provided by the Joint Typhoon Warning Center (JTWC) to evaluate risk based on a set of fixed parameters and limits. This risk assessment will lead to a recommendation by the Fleet Oceanographer to the Fleet Commander to set conditions of readiness (COR), or sortie assets to areas of safety. Destructive winds, storm surge, and sea heights are the most significant parameters upon which the Fleet Oceanographer bases his recommendation. However, it is destructive winds that carry the most weight when setting COR conditions. Guidance for what exactly defines the magnitude limit of

destructive winds, and the time period in advance for which action must be taken before these limits are reached, is stated in the instructions published by the Fleet, installation, and unit commanders.

While the Fleet Oceanographer has many sources from which to acquire TC forecast information, the official TC forecast for the WPAC is produced by the JTWC in Pearl Harbor, Hawaii. These official operational forecast products include written forecast discussions and visual decision aids that provide specific information regarding TC formation, position, projected track, and maximum wind speeds. While the offered TC forecast consists of specific track and intensity values, the uncertainty about these values can vary significantly. Many factors affect the uncertainty about a specific forecast. One indication of uncertainty is the degree of variability that may exist in the aids available to a forecaster. A wide range of forecast scenarios produced by independent numerical forecast models is typically an indication that uncertainty is high. While variability may change from forecast to forecast, current operational conveyance of uncertainty is based on a seasonally static, isotropic swath about the operational forecast.

One method used by JTWC to provide aid in generating forecasts and products is to construct a consensus of multiple independent operational deterministic forecast models. Goerss (2000) examined consensus forecasts to show that they typically have less error than any one single forecast. However, the issue of uncertainty was not addressed until Goerss (2007) used a regression analysis to show that the variability among deterministic forecasts could be used to define a measure of forecast uncertainty.

An alternative method to define uncertainty is to use an Ensemble Prediction System (EPS), which uses multiple perturbations of the initial conditions to an operational numerical weather prediction model. Based on the results of Goerss (2007), it may be expected that on average, the mean of individual ensemble forecasts would be more accurate than individual forecasts. Furthermore, a collection of individual ensembles prediction systems could be constructed to define a grand ensemble. Although

the ensemble-mean forecast may be more accurate than any one forecast, conveyance of uncertainty requires some measure of variability as defined by the individual ensemble member.

In this thesis, three individual EPS from three operational models are evaluated to examine the use of dynamic probability ellipses to better define the uncertainty in the prediction of a TC track. The relative skill of each ensemble-mean forecast TC tracks will be examined. These ensemble-mean forecasts are not compared to operational forecasts or deterministic forecasts due to differences in forecast reception relative to operational constraints for production of a forecast. Dynamic probability ellipses will be used to examine the character of each ensemble prediction system in relation to proper conveyance of uncertainty via examination of forecast reliability and resolution.

The products generated by JTWC, such as the official forecast track, wind speed probability swaths, and area of uncertainty, are weighed heavily upon by the Fleet Oceanographer when assessing risk. Operational JTWC forecasting methods, official TC track forecasts, and the production of the “area of uncertainty” will be further explained in Chapter II. The case of Super Typhoon Nanmadol (2011) demonstrates how challenging it can be to correctly forecast TC tracks when there is a high degree of uncertainty contained in numerical model forecast aids. This uncertainty makes it difficult for forecasters to determine forecast positions and track directions that are needed to better inform the Fleet Oceanographer and Commanders in their decision process. In the case of Super Typhoon Nanmadol, it was originally forecasted to impact Kadena Air Base, Okinawa, and possibly move to mainland Japan. It is probable that this gave base commanders and the 7th Fleet Oceanographer reasons to be concerned. Comparison of the official JTWC track forecasts (Figure 1) to the actual track, over the duration of the typhoon, conveys the difficulty in forecasting the motion of the storm. A brief synopsis of the storm behavior and effects by JTWC was entered into the Annual Tropical Cyclone Report (JTWC 2012a):

Super Typhoon Nanmadol (14W) formed within the monsoon trough east of the Philippines and began tracking west-northwestward toward Luzon in a complex steering environment dominated by a subtropical ridge to the north and east. It then took a poleward turn around the steering ridge,

rapidly intensifying to reach super typhoon status, and clipped the northeast tip of Luzon and then moved across the southern coast of Taiwan before dissipating in the Taiwan Strait, just prior to making landfall in China's Fujian Province. The cyclone reportedly caused at least 35 deaths and \$34.5M damage in the Philippines, at least 1 death and \$500M damage in Taiwan, and 2 deaths and \$48.5M damage in China.

While Nanmadol intensified to an estimated maximum wind speed of 140 kt and caused significant loss of life and property damage, the cyclone is also noteworthy from a forecaster's perspective as the numerical models and JTWC track forecasts depiction of an erroneous northeastward turn well to the east of the area eventually impacted by the cyclone. This tendency in both the model and subjective forecasts began during the cyclone development stage in the Philippine Sea and lasted well into its mature stage in the Luzon Strait. (JTWC 2012a)

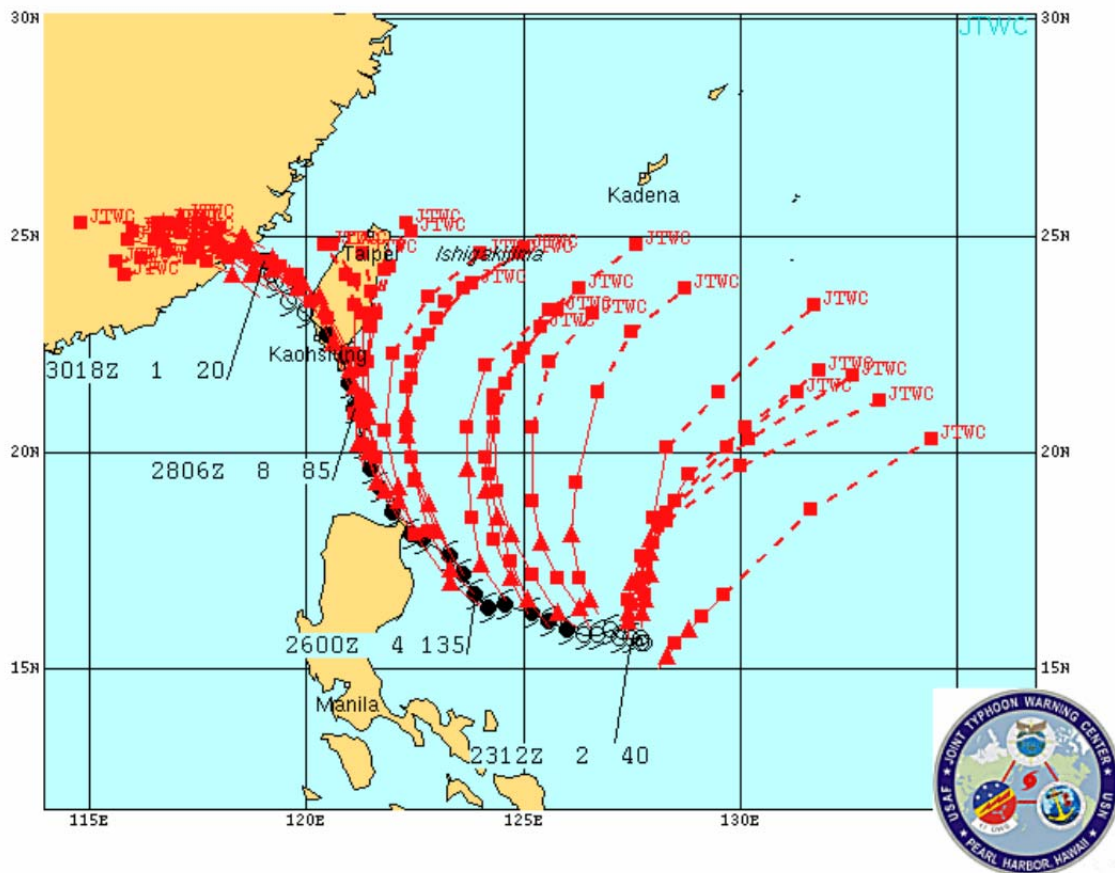


Figure 1. The JTWC track forecasts (squares) and best-track (tropical cyclone symbol) for STY 14W (From: JTWC 2012a)



## **B. OBJECTIVE**

Operational TC track forecasts provided by fleet weather centers require forecasters to assimilate guidance from operational numerical forecast models. As a result, the official forecast is often based on a subjective consensus of model guidance. The consensus mean of the deterministic models is often more skillful than individual operational forecasting systems during periods of increased uncertainty (Goerss 2007). An example of large forecast uncertainty is when TC track predictions have a large degree of variability among successive forecast integration. Several methods have been proposed to reduce uncertainty. The goal of this research is to analyze three operational EPS as an aid in identification of uncertainty in TC track forecasts. This will be accomplished by comparing the statistical characteristic of each EPS.

Recently, the ability of European Center for Medium-Range Weather Prediction (ECMWF) to predict that a TC would fall within a mean probability circle was studied by Majumdar and Finocchio (2009). It was found that the ECMWF ensemble mean (EEM) performed comparably to operational consensus techniques based on the mean of deterministic forecasts. Additionally, the study concluded that the EEM probability circles often did not contain the best-track position during periods of recurvature. Skill of forecasts defined by the EEM notably decreased in both EEM and in consensus-mean forecasts in the North Pacific during recurvature. The reason was primarily attributed to the fact that recurving TCs are often outliers of the ensemble and missed by the mean probability circles.

The operational weather prediction systems to be used in this study include the ECMWF, the United Kingdom Meteorological Office (UKMO) Unified Model, and the Japan Meteorological Agency (JMA) Global Spectral Model (GSM). While examining ensemble-based forecasts of TC tracks over the Atlantic Ocean, Nixon (2012) found that the ECMWF provided reliable forecasts in that the best-track positions occurred in the 68% probability ellipse of ensemble forecasts with a frequency of approximately 68%. The GFS and UKMO ensembles were found to be less reliable than the ECMWF. This

study will provide additional insight into the utility of EPS in providing measures of uncertainty of TC track predictions as no studies of this nature have been applied to the WPAC region.

The primary thrust of this research is to develop a spatial and temporal representation of the multiple single-model ensembles and demonstrate their utility in TC track prediction. Forecast tracks for the year 2009–2011 will be analyzed. These data sets will be obtained from the THORPEX Interactive Grand Global Ensemble (TIGGE), which will further be defined in Chapter III. It is expected that an appropriate representation of ensemble-based forecasts will provide added guidance that will reduce the forecast variability and assist in quantifying uncertainty in sortie and COR-setting decisions.

Background material is provided in Chapter II. The methodology used is described in Chapter III. The results are presented in Chapter IV and a case study and conclusions are given in Chapter V.

## **II. BACKGROUND**

### **A. JOINT TYPHOON WARNING CENTER OPERATIONAL METHODS TO DEFINE TROPICAL CYCLONE TRACK FORECAST UNCERTAINTY**

Operational TC track forecasts can be represented by three depictions of forecast uncertainty. Each depiction produced represents an increase in statistical representation of uncertainty. The operational track forecast is depicted in an official warning graphic (Figure 2) that defines the path of the storm center and an estimate of forecast uncertainty that is only based on historical errors. A wind speed probability swath (Figure 3) is defined based on a statistical sample of potential track errors by taking a random sample from the distribution of official track errors over the previous five year sample. Finally, a spread of deterministic models (Figure 4) is used to define the consensus that is used in conjunction with the variability among deterministic models to define forecast uncertainty.

#### **1. Official Forecast Track with Area of Uncertainty**

The JTWC predicted TC tracks are displayed on a two-dimensional forecast watch/warning graphic every six hours. The official forecast track is shown on this graphic as a pink line that defines the predicted track of the center of the TC through 120 h (Figure 2). The area of uncertainty is the shaded area around the forecast track, which is calculated by adding the JTWC 5-year running mean forecast track error to the forecast 34 knot wind radii at each forecast time. Since JTWC does not forecast wind radii at the 96- and 120-h, the area of uncertainty is calculated by adding the 72-hour 34 knot radii to the forecast track error at those times (JTWC 2012b). Thus, the shading highlights the area that may be affected by wind speeds exceeding 34 knots for a given JTWC forecast based on historical track forecast errors. However, this calculation does not account for uncertainty in the track forecast that may be based on an individual scenario or variation of numerical aids that are used as forecast guidance at a particular forecast time (JTWC 2012b).

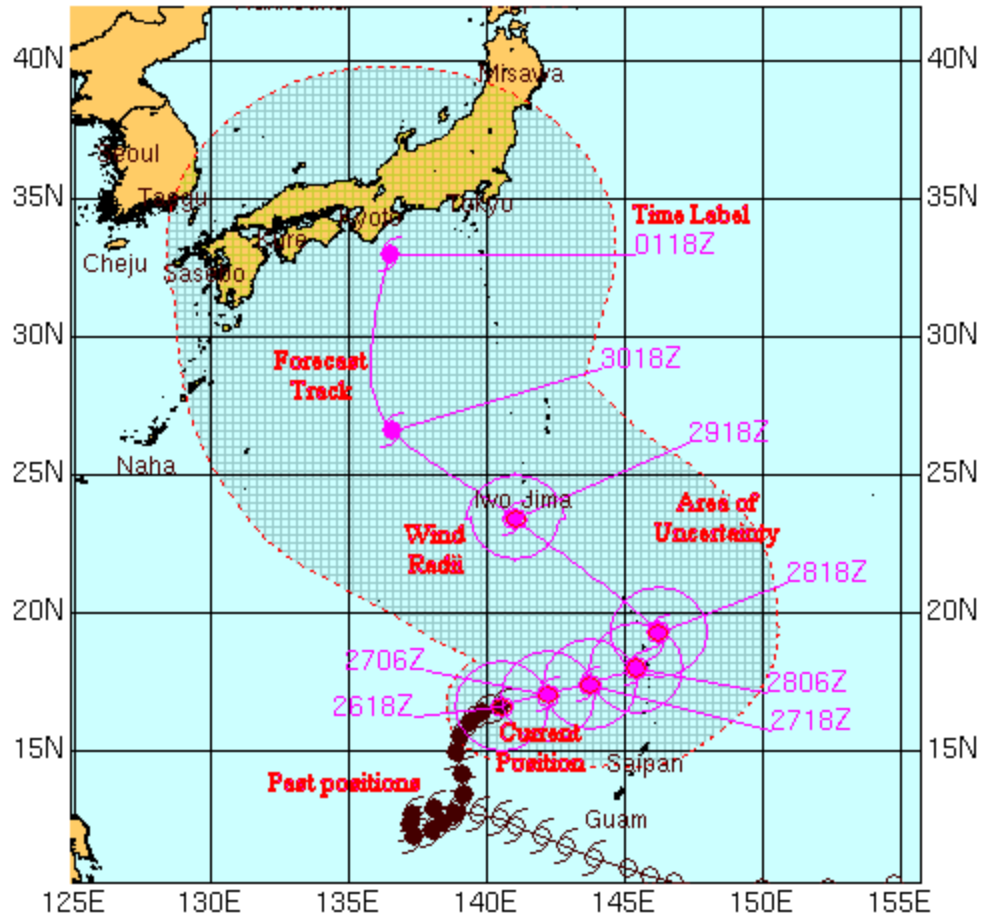


Figure 2. The JTWC Warning Graphic Legend. The solid black TC symbol defines the location of a TC, with winds greater than 64 knots. The last solid black TC symbol is the current position, and the pink cyclone symbols are forecasted positions. The pink circle, around the pink cyclone symbol, is the 64 knot wind radii. The shaded area of uncertainty is the 34 knot wind radii plus the average forecast track error (From: JTWC 2012b)

## 2. Wind Speed Probability Swath

Since 2006, the cumulative wind speed probability swath has been used as a tool for the Fleet Oceanographer and Commanders to better mitigate the risk to safety and damages due to winds. The JTWC produces a graphical display of the wind speed probability swath (Figure 3). The graphic can be viewed as 34, 50, or 64 kt cumulative probability wind speed swaths. It is produced every 12 h, over a 120-h forecast period. It is created by using a Monte Carlo Probability (MCP) model to produce 1000 TC track realizations (DeMaria et al. 2009). A single track realization is produced by taking a

random sample from the distribution of official track errors over the previous five year sample (DeMaria et al. 2009). All track error realizations are added to the official deterministic forecast tracks and then assigned an intensity and wind structure based on a wind profile model. A linear model is applied to account for serial correlation, track, and intensity dependency (DeMaria et al. 2009). The two major limitations of this product are that the sampling distributions do not account for any background flow dependences and sampling distributions are static for an entire TC season since they are based on the previous five hurricane seasons (Nixon 2012).

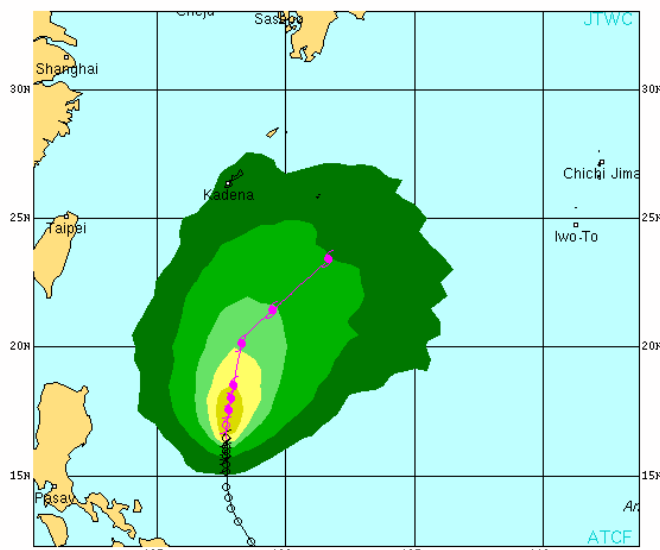


Figure 3. Wind speed probability graphic that depicts the likelihood of 50-kt winds will occur during the next 120 h issued 24 August 2011. Legend at the top of the box explains color scale representing the probabilities of 50-kt winds. The last black tropical cyclone symbol represents the position of the storm when the graphic was created (From: JTWC 2012c)

The GPCE is used as an aid in determining TC forecast tracks over the WPAC by predicting track error based on the spread of deterministic models that are used to define the consensus, which was found to be positively correlated with consensus model TC-track forecast error (Goerss 2007). The GPCE provides forecasters with a circle based on 70% confidence (Figure 4) that the analyzed TC will be located around a point defined

by the consensus of deterministic operational numerical models, which is labeled consensus version U (CONU). The CONU is computed when track forecasts from at least two of the following five models are available (Goerss 2007): Geophysical Fluid Dynamics Laboratory (GFDL) Interpolated (GFDI), Global Forecast System Aviation Interpolated (AVNI), Navy Operational Global Atmospheric Prediction System (NOGAPS) Interpolated (NGPI), the Navy GFDL Interpolated (GFNI), and the United Kingdom Met (UKMET) Office Global Model Interpolated (UKMI). Interpolated models are defined by moving the model forecast position in space to account for time offset between the model initialization time and forecast time. The GPCE technique is based on the assumption of an isotropic error distribution around the consensus mean. Based on regression analysis between the spread of the deterministic model and forecast accuracy, Goerss (2007) defined the isotropic circle size can be defined such that the probability of best-tracks is within the circle is 70%.

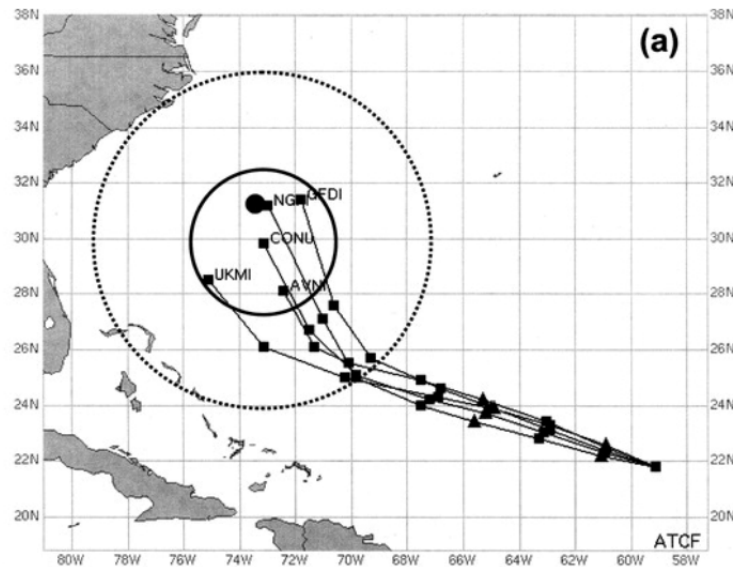


Figure 4. Predicted 70% confidence radius (solid circle) of the 120-h CONU forecast for Hurricane Isabel on 0000 UTC 13 September 2003. The individual model tracks used to create the CONU consensus model are shown. Notice the GPCE circle is much smaller than the 120-h radius (dotted circle) used by the NHC potential 5-day track area graphic (From: Goerss 2007)

#### 4. Goerss Predicted Consensus Error Along- and Across-track (GPCE-X)

The GPCE-X is a modification to GPCE that removes the requirement for an isotropic distribution about the consensus mean (Hansen et al. 2010). In GPCE-X anisotropic ellipses are defined based on a partitioning of the forecast track error into components that are along-track and cross-track (Figure 5). The ellipse is defined to contain 70% of the predicted positions of the members contained in the consensus of deterministic models (CONU).

The XTE defines the portion of the predicted error in the consensus track that is to the left or right of the verifying position. The ATE defines the portion of the predicted error of the consensus track that is ahead and behind the consensus mean position. The GPCE-X improves upon the sharpness of the predicted area by considering along and across-track spread, which reduces the area of the ellipse (Pearman 2011).

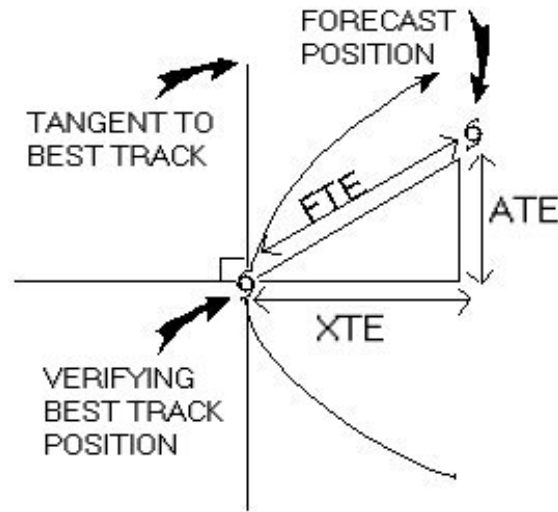


Figure 5. Schematic definition of forecast-track error (FTE), cross-track error (XTE), and along-track error (ATE) (From: JTWC 2012a)

#### B. ENSEMBLE PREDICTION SYSTEMS (EPS)

For analysis of TC tracks, there are a variety of EPS from which to choose. Pearman (2011) and Nixon (2012) focused their studies on the Atlantic basin, and provided thorough descriptions of ECMWF, UKMET, and National Centers for Environmental Prediction (NCEP) EPS in their research.

In this study, the addition of JMA/GSM data is important and appropriate due to the regional use of the JMA forecasts. The NCEP ensemble will not be used in this thesis due to the lack of ensemble track forecasts in the JTWC data archive.

The ECMWF, UKMET, and JMA employ different methods to create the perturbations needed to generate ensemble members. The ECMWF and JMA both create initial perturbations based on the singular vector method, and the perturbations are added to a control analysis. The singular vectors with large singular values represent fast-growing perturbations over a prescribed time interval under the assumption that the perturbations grow linearly (Lorenz 1965). The fast-growing perturbations are considered to be responsible for large forecast uncertainty at the optimization time leading to sufficient dispersion in the most uncertain directions (Yamaguchi and Majumdar 2010). The UKMET uses an Ensemble Transform Kalman Filter (ETKF) to create initial perturbations, which are then added to a control analysis. An ETKF provides estimates of the true state of the atmosphere that is based upon an optimized blend of short-term forecasts and current observations (Bowler et al. 2008). The optimization depends on statistical characteristics defined by the ensemble members associated with the data assimilation.

In general, the purpose of ensemble model forecasting is to define a flow-dependent representation of model forecast uncertainty that is based on the overall chaotic nature of the atmosphere and its representation by the governing equations of motion (Lorenz 1963). The first two sections provide a description of the EPS from ECMWF and UKMET based on descriptions of Pearman (2011) and Nixon (2012). The third section provides a description of the JMA/GSM EPS.

## **1. European Center for Medium Range Weather Forecasts (ECMWF)**

The ECMWF EPS has a horizontal resolution of 32 km. There are 50 ensemble members and one control member for a total of 51 ensemble members. The control member is defined as the deterministic forecast system with no perturbation except



integrated at coarser resolution than the operational deterministic model. The EPS forecasts are initialized every 12 h at 0000 UTC and 1200 UTC. The output forecasts extend out to 384 h at an interval of 12 h.

The 50 perturbations are created by a combination of three methods: (1) singular vector (SV) technique; (2) using differences between the members of an ensemble of data assimilation schemes (EDA); and (3) using two different stochastic perturbation techniques (ECMWF 2012). The SVs are selected based on the greatest linear growth rate in total energy over a 48-h time period for a fixed set of norms, assumptions, and spatial targets (ECMWF 2012). Using differences between the members of EDAs is accomplished by utilizing a set of 6-hour forecasts starting from ten different analyses that differ by means of small variations to the observations and the stochastic physics (ECMWF 2012). The two different stochastic perturbation techniques are stochastic physics and stochastic backscatter. In stochastic physics, the tendencies in the physical parameterization schemes are randomly perturbed. In the stochastic backscatter model perturbations are added to the vorticity tendencies to replace the kinetic energy damped in parameterization that account for unresolved scales (ECMWF 2012). Each method yields a different set of perturbations over the Northern and Southern Hemispheres, and the tropics. These perturbations are combined linearly, multiplied by coefficients randomly sampled from a Gaussian distribution, and then mirrored by reversing the signs of the 25 members created to yield a total of 50 global perturbations (ECMWF 2012).

## **2. United Kingdom Meteorological Office (UKMO)**

The UKMO EPS is called the Met Office Global and Regional Ensemble Prediction System (MOGREPS) and has horizontal resolution of 60 km and consists of 23 ensemble members constructed from 22 perturbation members created by the use of an ETKF, and one control. The MOGREPS is initialized every 6 h at 0000 UTC, 0600 UTC, 1200 UTC, and 1800 UTC. The output forecasts are available to 144 h at an interval of 12 h.

### **3. Japan Meteorological Agency Global Spectral Model (JMA/GSM)**

The JMA EPS is based on the GSM with a spatial resolution of 20 km and has 51 members defined by 50 ensemble members and one control. The EPS is initialized every 12 hours at 0000 UTC and 1200 UTC. The output forecasts are available every 6 h out to 216 h. The JMA EPS is similar in construct to the ECMWF EPS in that the SV technique is used to define the initial perturbation (ECMWF 2012).

### **III. METHODOLOGY**

#### **A. DATA**

##### **1. Data Source**

The data for this study include all forecast and actual TC positions for the WPAC from 2009-2011. A total of 54 named TCs formed over this period, of which 20 were tropical storms (34 to 63 kt winds), 24 were typhoons (64 to 130 kt winds), and 10 were super typhoons (greater than 130 kt). The total number of storms in each year was below the climatological mean, which is 31 TCs per year. The most significantly below average year was 2010 with only 14 named storms. Data for all three years were available and used to determine the statistics and tendencies of the EPS forecasts.

The 2009 typhoon season (Figures 6a and 6b) was the most active of the three years, with 22 named storms consisting of seven tropical storms, ten typhoons, and five super typhoons. There were 13 typhoons that made landfall, or came close to island chains in various regions; nine in the South China Sea region over the Philippines and Southeast Asia mainland, two near Guam, and two near Okinawa. The 2009 JTWC Annual Tropical Cyclone Report best summarizes the major storms for the season:

Official and media reports indicated that Typhoon 09W (Morakot) was the most destructive cyclone in the western North Pacific in 2009, with significant damage reported on Taiwan and Fujian, Zhejiang and Jiangsu provinces, China as well as approximately 650+ fatalities reported mainly due to landslides from extremely heavy rainfall. STY 20W (Melor) was the only system to make landfall near DoD installations in 2009. This cyclone formed to the east of Guam, crossed the northern Marianas Islands and attained super typhoon intensity to the southeast of Okinawa Island. STY 20W subsequently went ashore about 150 miles southwest of several DoD installations in the Kanto Plain as a minimal typhoon. (JTWC 2012a)

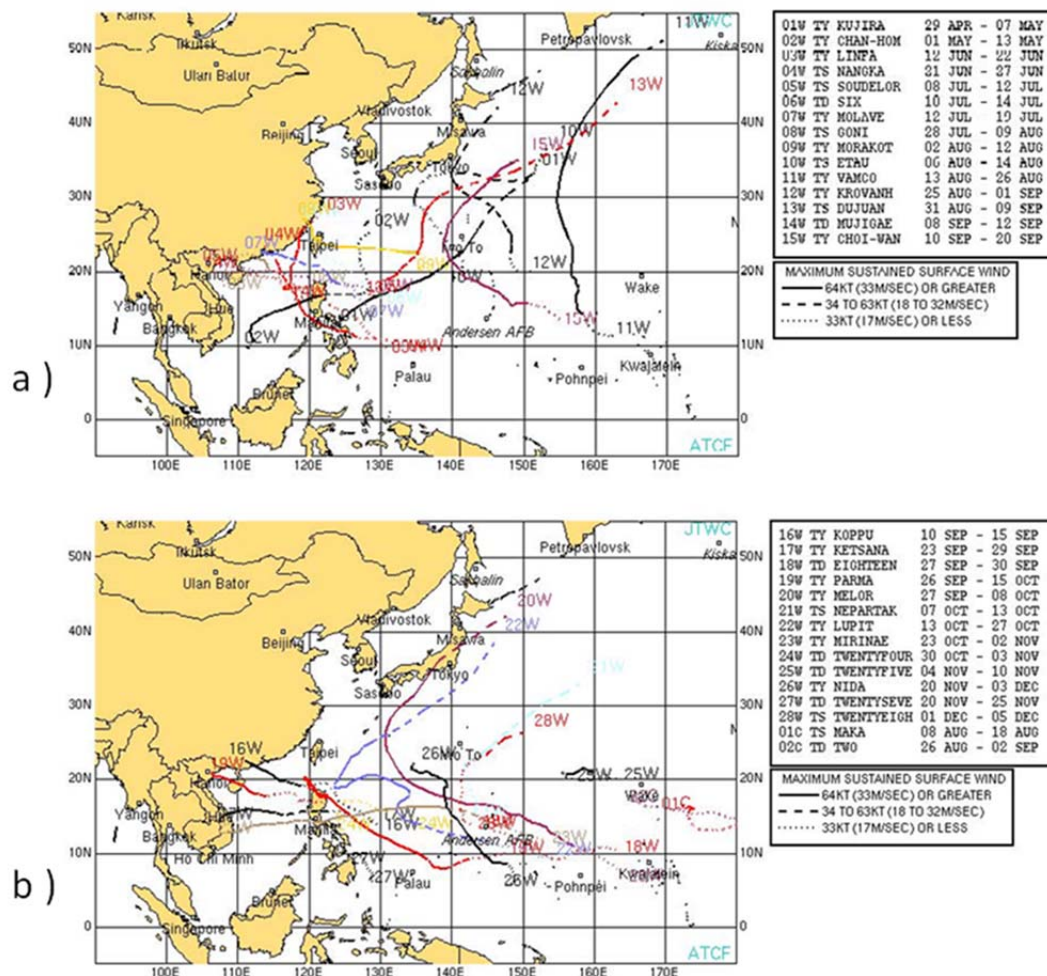


Figure 6. Tropical cyclones that occurred during (a) April through early September 2009, and (b) September through December 2009 (From: JTWC 2012a)

The 2010 typhoon season (Figures 7a and 7b) was the least active of the three years, with 14 named storms that included five tropical storms, eight typhoons, and one super typhoon. There were 11 typhoons that made landfall, or came close to island chains in various regions; seven in the South China Sea region over the Philippines and Southeast Asia mainland, three near Okinawa with one TC that continued north to make landfall over the Korean Peninsula. The 2010 JTWC Annual Tropical Cyclone Report best summarizes the major storms for the season:

Super Typhoon Megi (15W) was the only cyclone to reach super typhoon intensity. Typhoon Kompasu (08W) was operationally significant because it made landfall on Okinawa as a strong typhoon and on South Korea as a minimal typhoon. (JTWC 2012a)

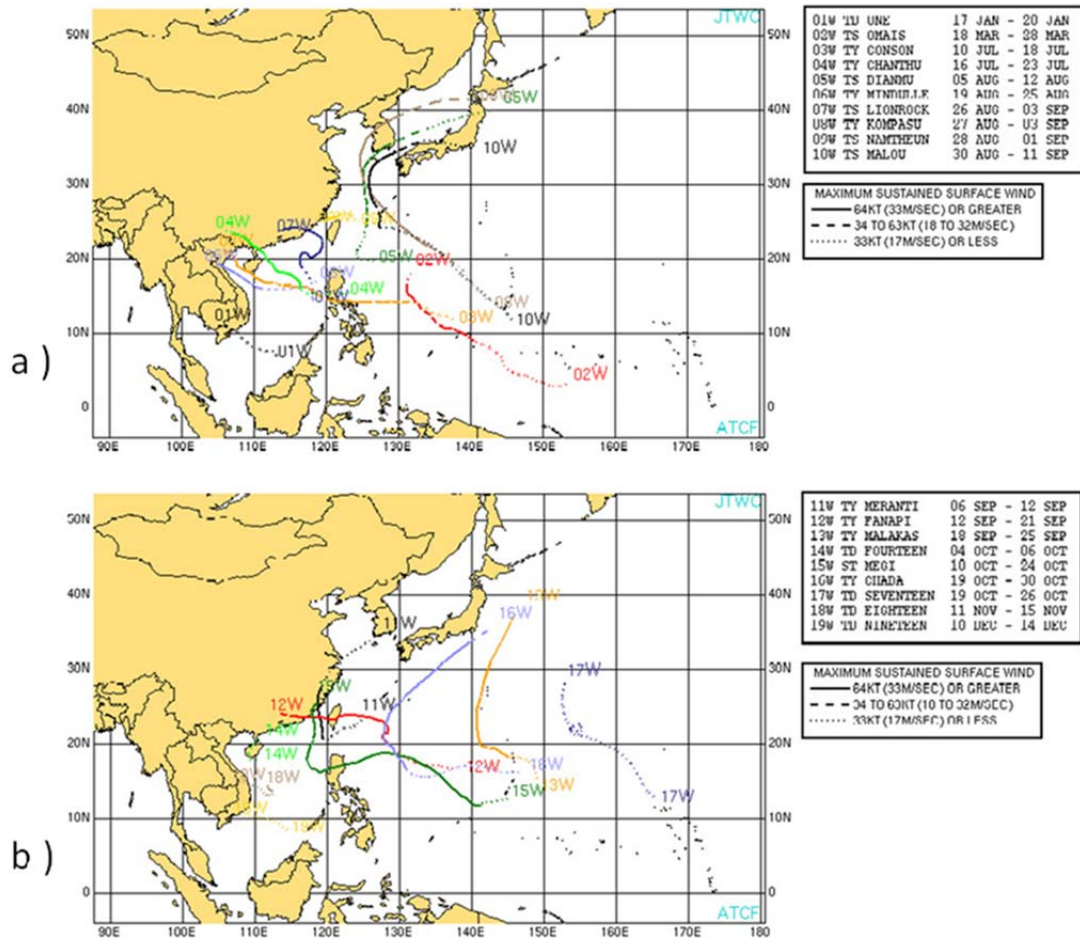


Figure 7. Tropical cyclones that occurred during (a) April through early September 2010, and (b) September through December 2010 (From: JTWC 2012a)

The 2011 typhoon season (Figures 8a and 8b) had 18 named storms consisting of eight tropical storms, six typhoons, and four super typhoons. There were nine typhoons that made landfall, or came close to island chains in various regions; four in the South China Sea region over the Philippines and Southeast Asia mainland, two over southern Japan, and three near Okinawa. The 2011 JTWC Annual Tropical Cyclone Report best summarizes the major storms for the season:

The TC formation region was displaced north and west again in 2011, a characteristic common during La Nina conditions. Several of these early to mid-season forming TCs exhibited S shaped, looping, or generally erratic tracks, with numerous passages near or over Okinawa. In fact, Super Typhoon Songda (04W) passed just west of Kadena Air Base and destroyed the WSR-88D Doppler Weather Radar. (JTWC 2012a)

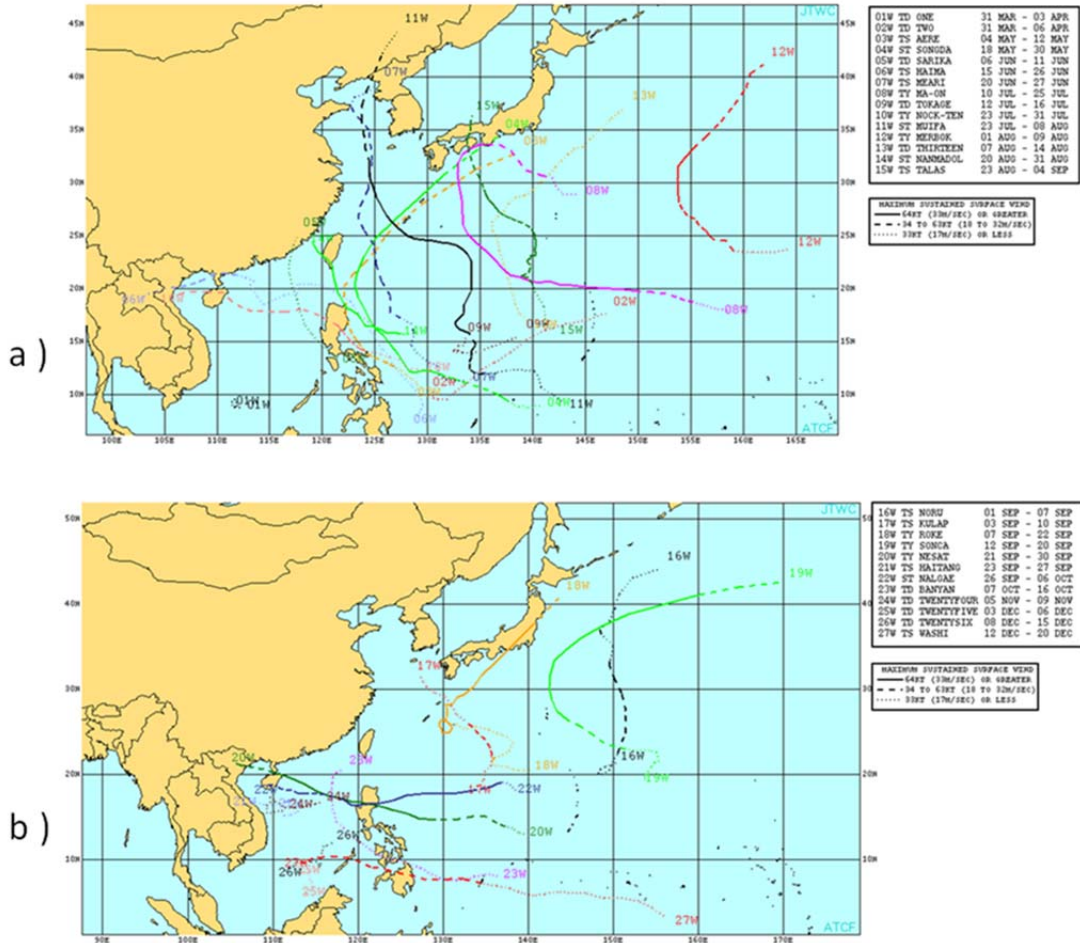


Figure 8. Tropical cyclones that occurred during (a) April through early September 2011, and (b) September through December 2011 (From: JTWC 2012a)

## 2. Data Format

The data used in this study were produced by three operational EPS that are contained in the TIGGE (THORPEX Interactive Grand Global Ensemble) database online at <http://tigge.ucar.edu/home/home.htm>. The standard data format is in Cyclone XML (CXML), which is a descriptive, human-legible format used to define forecast track

positions for all global EPS (Pearman 2011). The CXML format is defined such that it is comprised of data from multiple simultaneous cyclone observations, analyses, deterministic and ensemble numerical model forecasts (AGBOM 2012). Tropical cyclone forecasts are placed in the TIGGE CXML data archive for all storms that are at least tropical storm intensity. The storm positions are defined for each EPS based on individual definitions of a low level circulation center.

Best-track data are defined by JTWC, and shared online at [http://www.usno.navy.mil/NOOC/nmfc-ph/RSS/jtwc/best\\_tracks/](http://www.usno.navy.mil/NOOC/nmfc-ph/RSS/jtwc/best_tracks/). The best-track storm position and intensity are defined by post-season analysis, which includes surface observations, satellite images, aircraft reports, and radar images. A best-track entry exists for every six hour period (0000 UTC, 0600 UTC, 1200 UTC, 1800 UTC) over the lifetime of the storm (Nixon 2011).

### **3. Data Homogeneity**

The forecast track datasets were compiled and analyzed for individual EPSs, the WPAC region, and specific sub-regions (Figure 9). Due to model differences, not all models will contain forecasts of each storm at every time. Some models may form a TC later or earlier than others. Therefore, a homogeneous dataset is defined such that only cases that have forecasts for all three EPS are included. In addition, any TC that did not persist longer than 48 hours was removed from the dataset. An additional homogeneity constraint was applied for the GE, GPCE and GPCE-X forecasts.

### **4. Region and Sub-regions**

The database used in this study was compiled specifically for the WPAC region, and then divided into specific sub-regions to examine detailed variabilities in forecast accuracy and uncertainty. The Monsoon Trough Region (MTR) defines the mean position of the WPAC monsoon that provides favorable low-level, large-scale cyclonic shear in which convective disturbances form. Most of the TCs that form in the MTR move westward to the South China Sea Region (SCSR) or recurve poleward. Recurvature typically occurs when the subtropical ridge is weakened and the storm moves northwestward due to the weakened easterly winds on the equatorward side of the



subtropical ridge. The SCSR is used to capture TCs that move westward under influence of the subtropical ridge. The Recurve Region (RCR) defines the latitudinal region where TCs normally recurve. Following recurvature, TCs move into the Mid-Latitude Region (MLR) and typically move westward under the influence of mid-latitude westerlies. The number of TC passing through each region in Figure 9 is defined in Table 1.

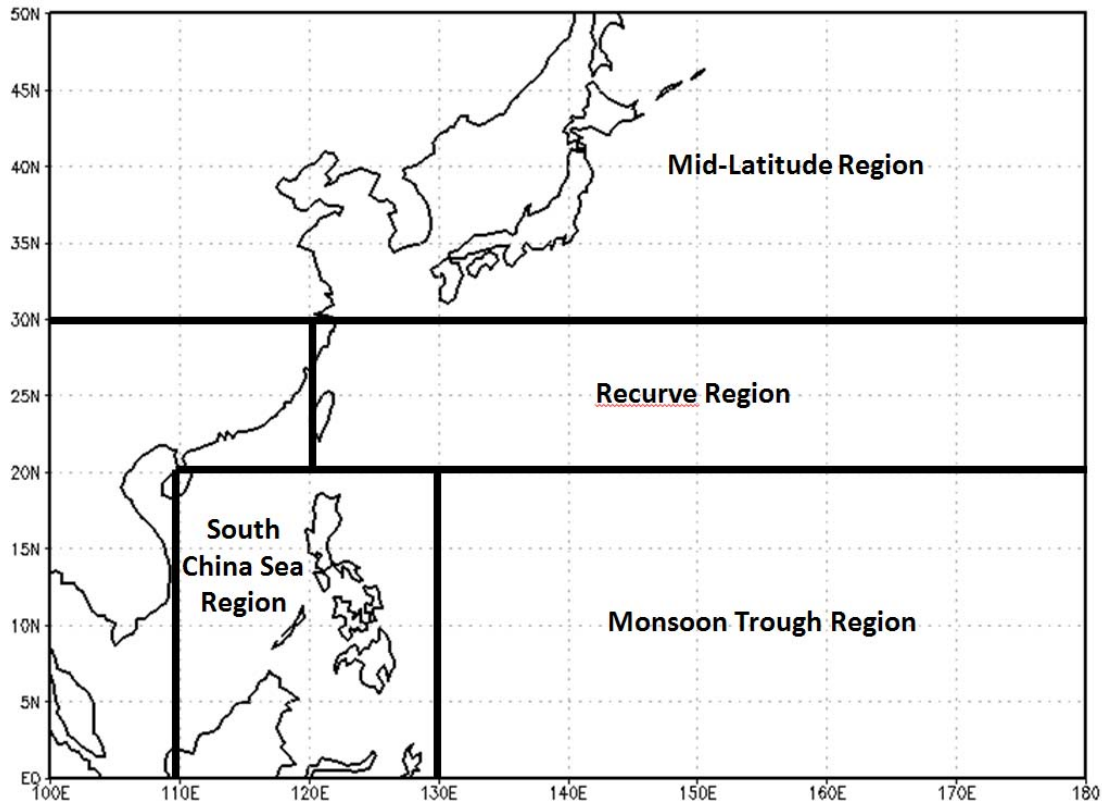


Figure 9. Geographic sub-regions of the WPAC region used to group Ensemble Prediction System (EPS) forecast track data

Table 1. Total number of Tropical Cyclones included in this thesis for the WPAC and each sub-region

Year	Pacific Basin TCs	Monsoon Trough Region TCs	South China Sea TCs	Recurve Region TCs	Mid-Latitude Region TCs
2009	22	9	12	11	4
2010	14	5	11	6	5
2011	18	7	6	11	9
<b>TOTAL:</b>	54	21	29	28	18



## 5. Developing the EPS Ellipse

As explained in Chapter II (Figure 2), the creation of JTWC forecast track and area of uncertainty is one of a number of methods used to analyze and display forecast uncertainty. The methods have limitations due to the lack of confidence for TC position inside of the area of uncertainty. Carr and Elsberry (2000) quantified the relationship of the spread in the consensus of deterministic models to forecast track error. Additionally, Pearman (2011) concentrated on understanding the relationships and disadvantages between the current methods of calculating uncertainty by creating a Grand Ensemble (GE) forecast track ellipse and comparing EPS, GE, GPCE, and GPCE-X forecasts.

Pearman (2011) examined the uncertainty in the GE forecast position as being defined in relation to the principal axis of the spatial distribution of EPS members centered relative to the GE mean position. The ellipse is defined to contain 68% of the GE member forecast track positions, is centered on the GE mean, and is calculated for each forecast period in the homogeneous dataset (Pearman 2011).

The ellipse calculation performed by Pearman (2011) is a critical procedure in understanding the ensemble member spatial distribution. This study employs the same procedure to create the EPS and GE ellipses. The initial step is to create a  $2 \times n$  matrix (**A**) of the latitudes and longitudes of each ensemble member contained in the EPS, at a particular forecast time. The variable  $n$  is defined as the total number of forecasts present. A covariance matrix is created from the matrix of latitudes and longitudes:

$$\Sigma_x = \sigma_0^2 (\mathbf{A}^T \mathbf{A})^{-1} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix} \quad (1)$$

The next step is to calculate eigenvalues ( $\lambda_1$ ,  $\lambda_2$ ) and eigenvectors of the covariance matrix:

$$\begin{Bmatrix} \lambda_1 \\ \lambda_2 \end{Bmatrix} = \frac{1}{2} (\sigma_1^2 + \sigma_2^2 \pm \sqrt{(\sigma_1^2 + \sigma_2^2)^2 - 4(\sigma_1^2 \sigma_2^2 - \sigma_{12}^2)}) \quad (2)$$

The resulting eigenvalues scale the ellipse along the semi-major and semi-minor axes, while the eigenvectors of the covariance matrix define the orientation, which orients of the ellipse axes to be in the direction of the largest variance.

The mean of the latitude and longitude matrix is computed to find the center position of the ellipse. To ensure that the ellipse captures 68% of the EPS member forecast positions, a Chi-squared distribution is assumed and a Chi-squared scaling parameter is applied.

Based on the Chi-squared scaling, the ellipses were defined such that an ellipse captures 68% of the EPS ensemble members at a 95% confidence interval, using a homogeneous dataset (Figure 10). The orientation of the semi-major axis is determined by the spatial distribution of the EPS members.

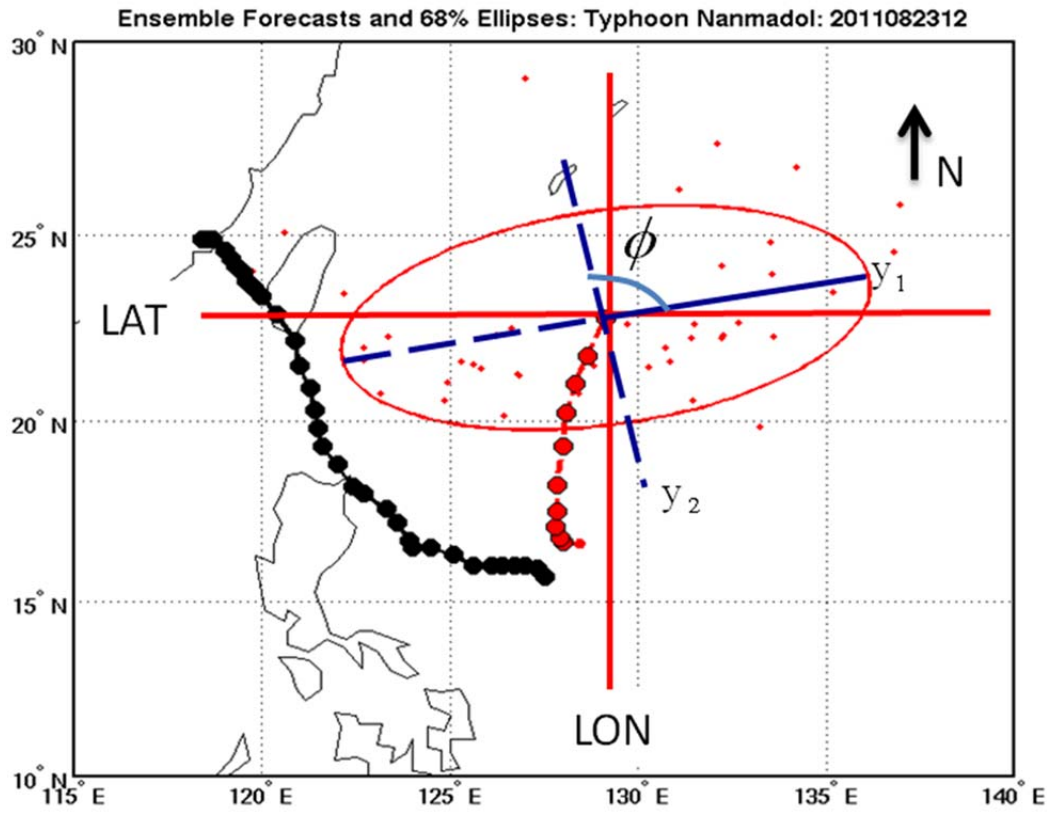


Figure 10. The probability ellipse (red) contains 68% of the ensemble members (red dots) and is centered on the ellipse mean position.  $y_1$  and  $y_2$  are the coordinate system by which the ensemble members are rotated around. The orientation of the ellipse is defined by  $\phi$

## B. STATISTICAL ANALYSES

Nixon (2011) addressed the statistical characteristics of ensemble forecasts by computing the reliability and resolution of individual EPS relative to the 68% ellipse resolution. Reliability of a forecast pertains to the calibration of the forecast, which relates the forecast to the observation. For forecasts that are perfectly reliable, the observed relative frequency of an observation would be equal to the forecast probability. For TC location prediction, we define reliability from 68% threshold, when the 68% probability ellipse encompasses the best-track position 68% of the time. The resolution of a forecast pertains to the ability of a forecast system to sort observed events into groups that are different than each other. If the observed position of a TC is quite different in forecasts of recurvature versus straight moving, then the forecast system has adequate resolution to identify these two scenarios. If the observed location is the same for a forecast of recurvature as for a forecast of straight moving, then the forecast system cannot resolve these two types of forecasts. The resolution and reliability of each forecast interval will be inferred jointly by calculating both the Probability within Spread (PWS) and Mean Area Difference (MAD).

### 1. Probability within Spread

Probability within Spread (PWS) defines the probability that an observed TC is within the spread of an EPS as:

$$\text{PWS}(k) = \frac{1}{M} \sum_{m=1}^M \left\{ \frac{0:s_{obs} > k(\sigma)_m}{1:s_{obs} \leq k(\sigma)_m} \right\} \quad (3)$$

where  $k, m$  are integers ( $k = 1, 2, 3 \dots$ ),  $M$  is the total number of forecasts at a given lead time,  $s_{obs}$  is the distance of the observed TC from the EPS mean, and  $\sigma$  is the across-track spread of the EPS (Nixon 2011). If members are sampled from a normal distribution, a PWS value of 0.68 for a spread defined by one standard deviation about the ensemble mean would indicate statistical consistency (Buckingham et al. 2010). Furthermore, if members are sampled from a normal distribution and the spread is  $1\sigma$  ( $k = 1$ ),  $2\sigma$  ( $k = 2$ ), and  $3\sigma$  ( $k = 3$ ), PWS values of 0.68, 0.95, and 0.99 would indicate statistical consistency (Buckingham et al. 2010). The spread is defined by the standard deviation of the EPS members about the ensemble-mean position. PWS calculates a percentage that reflects a

spatial representation of best track presence within a cross-track conically bounded area. For the purposes of this study, a PWS value of 0.68 for 1  $\sigma$  was utilized to indicate statistical consistency.

## 2. Ellipse Reliability

Ellipse reliability is the percentage of time that the best-track analysis position is within the EPS ellipse at a particular forecast time. Reliability is defined as:

$$\frac{1}{M} \sum_{m=1}^M \left\{ \frac{0:s_{obs} > (68\% \text{ ellipse})_m}{1:s_{obs} \leq (68\% \text{ ellipse})_m} \right\} \quad (4)$$

where the definition of the EPS ellipse is defined as enclosing 68% of the ensemble forecast track members, resulting in the expected ellipse reliability to be 68% (Nixon 2011). The reliability percentages of the EPS that are higher (lower) than 68% will be determined to be under (over) dispersive. The ellipse reliability statistical analysis method is similar to PWS in that it calculates a percentage of events in which the best-track is located in a defined area. However, PWS calculates a percentage that reflects a spatial representation of best track location in a cross-track conically bounded area. Ellipse reliability calculates a percentage that reflects a temporal representation best-track location within a defined ellipse that is defined to encompass a set number of ensemble members. Furthermore, the orientation of the ellipse carries based on flow-dependent uncertainty, which the speed is based on standard deviation about the mean will be symmetric by definition.

## 3. Mean Area Difference (MAD)

The Mean Area Difference (MAD) is a calculation that compares the area of the EPS ellipse, which is the forecast ellipse area, with a control ellipse area. This provides a percentage difference in area, or MAD, which is defined as:

$$MAD = \frac{Control \text{ Ellipse}_{area} - Forecast \text{ Ellipse}_{area}}{Control \text{ Ellipse}_{area}} \quad (5)$$

A MAD percentage is positive (negative) if the EPS ellipse area is less (more) than the control ellipse area.

## **IV. ANALYSIS AND RESULTS**

### **A. OVERVIEW**

Each individual EPS is characterized by an ensemble-mean track forecast and a spread or dispersion about the mean forecast. As defined in Chapter II, the primary statistical measures used to evaluate the ensemble-mean forecast are the FTE, XTE, and ATE. These measures are analyzed to determine whether any significant biases or systematic errors exist in the ensemble-mean forecast tracks. The statistical analyses are examined for the entire WPAC basin and for each sub-region.

### **B. ENSEMBLE-MEAN TRACK ERRORS**

#### **1. FTE**

The FTE is defined as the absolute distance between the ensemble-mean and the TC best-track positions. The FTE indicates which EPS forecast tracks are closest on average to the actual best-track analysis positions. The spread about the FTE will be indicated by vertically-oriented bars that define the standard deviation for each mean FTE. The standard deviation is defined as the square root of the average squared difference between the ensemble-mean track position and the TC best-track analysis position. Because the standard deviation involves squared errors, even one very large track error can strongly impact the magnitude of the standard deviation in the bar graph. Based on the FTE, no directionality information can be inferred with respect to the track error characteristics for each EPS.

For all forecasts over the entire WPAC (Figure 11), the ECMWF ensemble-mean FTE is smaller than the FTE of the UKMO and JMA ensemble-mean forecasts. Additionally, the standard deviation about the ECMWF FTE is smaller than the standard deviation about the mean FTE for the UKMO and GFS EPS.

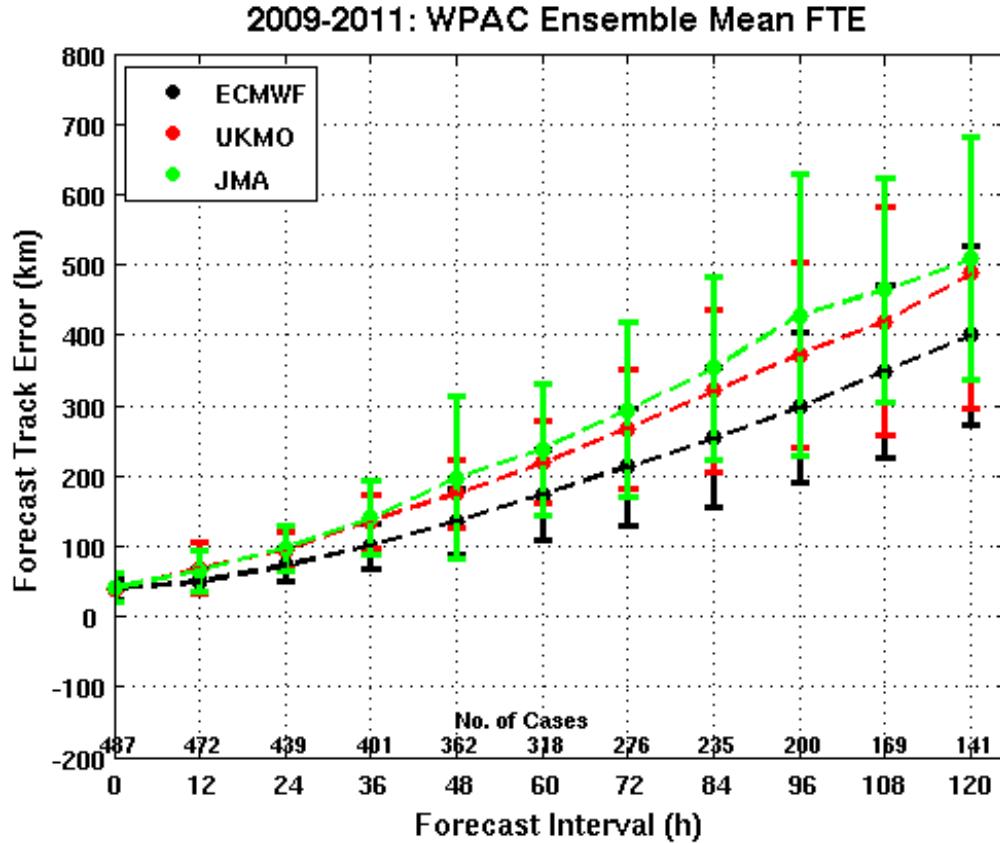


Figure 11. Average FTE for each EPS over the entire WPAC from 2009–2011. Plus and minus one standard deviation about the FTE is represented by a vertically oriented line at each 12-h forecast from 0–120 h

For the Monsoon Trough Region (MTR) (Figure 12), the FTE of the ECMWF ensemble-mean forecast is also smaller than that of the ensemble-mean forecasts from the UKMO and JMA. At forecast ranges between 48–84 h, the ensemble-mean FTE for the JMA and UKMO are approximately equal. For the ECMWF forecast intervals shorter than 72 h, the FTE over the MTR (Figure 12) is approximately equal to the FTE over the entire WPAC (Figure 11). Beyond 72 h, the ECMWF ensemble-mean FTE over the MTR is less than that over the entire WPAC. By 120 h, the ECMWF ensemble-mean FTE over the MTR is 100 km less than the WPAC average. The UKMO ensemble-mean FTE over the MTR is also less than the total WPAC value. The JMA ensemble-mean FTE over the MTR is approximately the same as the WPAC average.

From 0 to 96 h the FTE standard deviations for ECMWF are noticeably lower in the MTR than the values over the entire WPAC and only slightly lower at 108 and 120 h. The small FTE and standard deviation of the ECMWF ensemble-mean FTE over the MTR is probably due to the relatively straight steering flow over the MTR, which is typically less variable than the more complex steering flows over other regions of the WPAC.

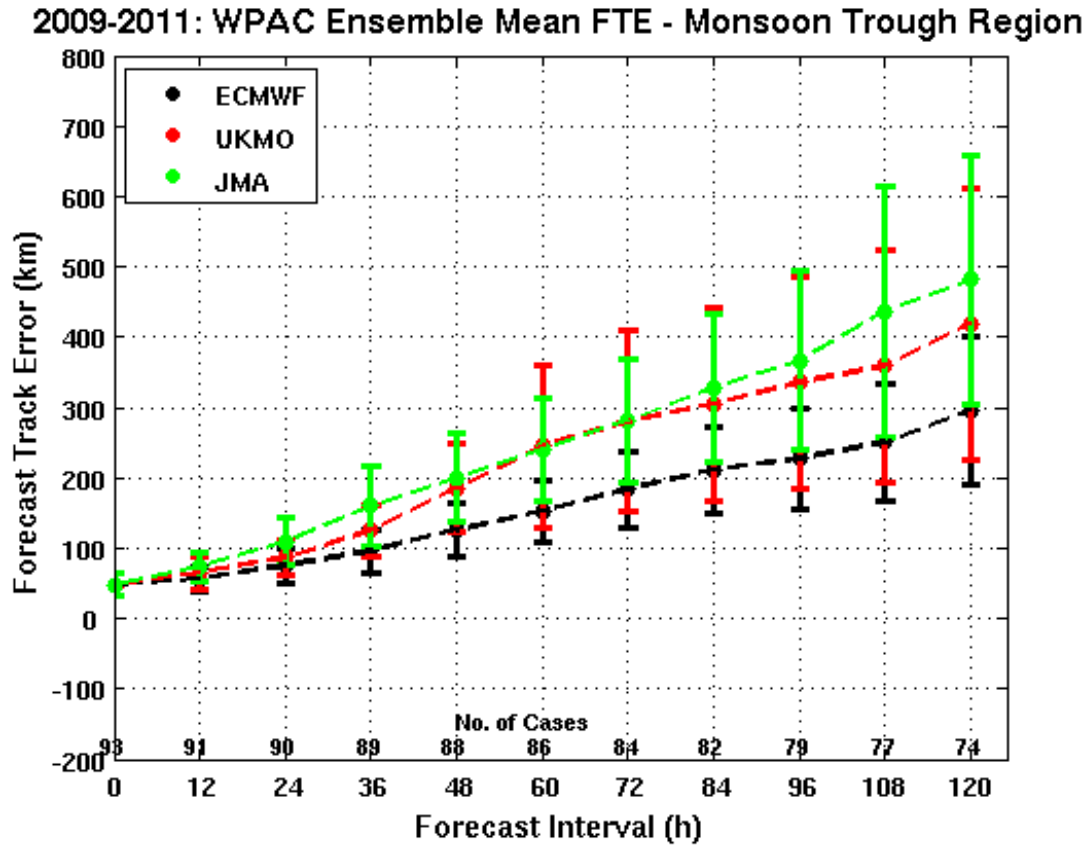


Figure 12. As in Figure 11, except for the MTR region

For the Recurve Region (RCR) (Figure 13), there is a noticeable increase in the ensemble-mean FTE for all EPS beyond 48 h. On average, the ensemble-mean FTE is 50-100 km larger in this region than in the total WPAC (Figure 11). The increase in FTE can be attributed to the more variable steering flow that is influenced by mid-latitude

flow patterns. Additionally, the location of recurvature is often highly sensitive to the interaction between the mid-latitude flow poleward of the TC and the subtropical ridge equatorward and east of the TC.

The standard deviation about the ensemble-mean FTE is greater for ECMWF in this region as compared to the WPAC and other sub-regions. Noticeably, UKMO has the largest mean FTE but smaller standard deviations when compared to the other sub-regions indicating it has a larger FTE bias in the RCR. The JMA ensemble-mean FTE errors in this region are comparable to the FTE in other regions and the overall WPAC.

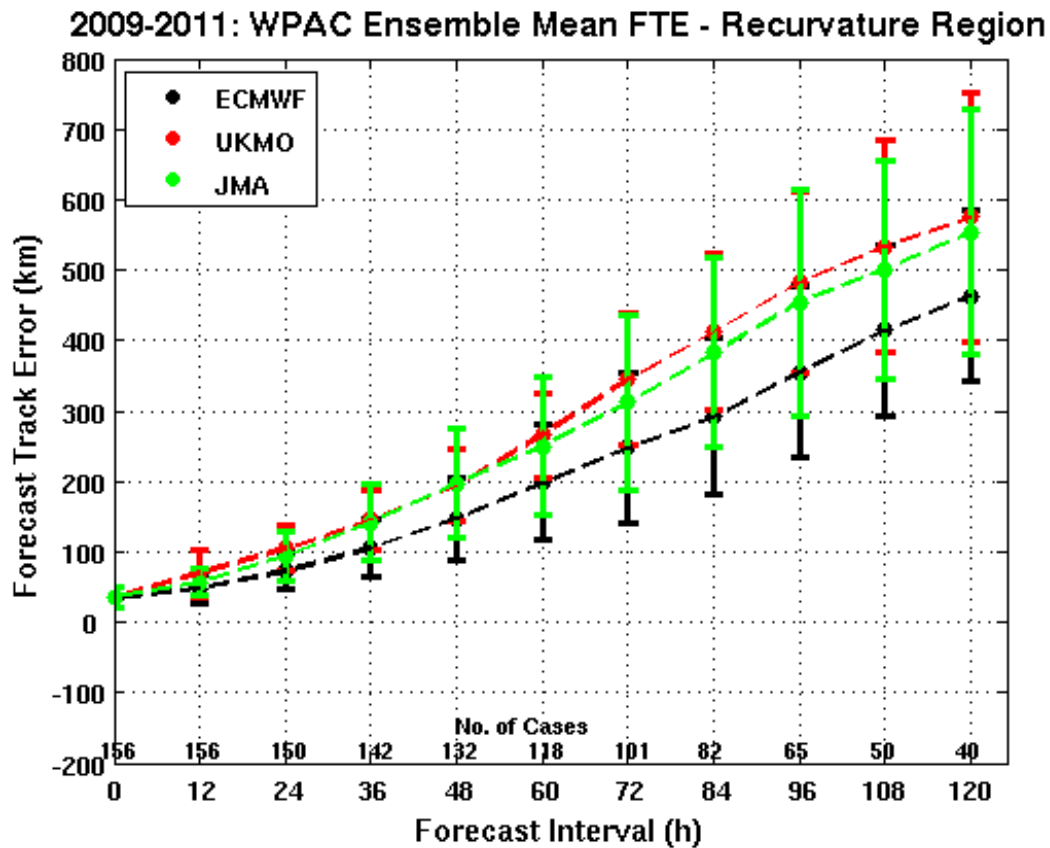


Figure 13. As in Figure 11, except for the RCR region

Over the South China Sea Region (SCSR) (Figure 14), the ensemble-mean FTE for the ECMWF are similar to the values observed over the entire WPAC (Figure 11). At forecast ranges from 12–96 h, the UKMO ensemble FTE over the SCSR is less than that over the entire WPAC. At forecast ranges between 48–108 h, the average FTE for JMA is 50 km larger than that over the WPAC.



The standard deviations about the ensemble-mean FTE for JMA from 48-120 h are larger compared to the standard deviations for the same forecast intervals over the entire WPAC. The increased FTE and larger standard deviations for the JMA could indicate poor forecasts of flow patterns over the SCSR, which could lead to early recurvature of the TCs in the mid-range forecasts. The UKMO values are lower with smaller standard deviations from 12-96 h when compared to that of the WPAC. Therefore, the UKMO model forecasts are more accurate over the SCSR. The ECMWF FTE over the SCSR are just slightly larger than the FTE over the entire basin.

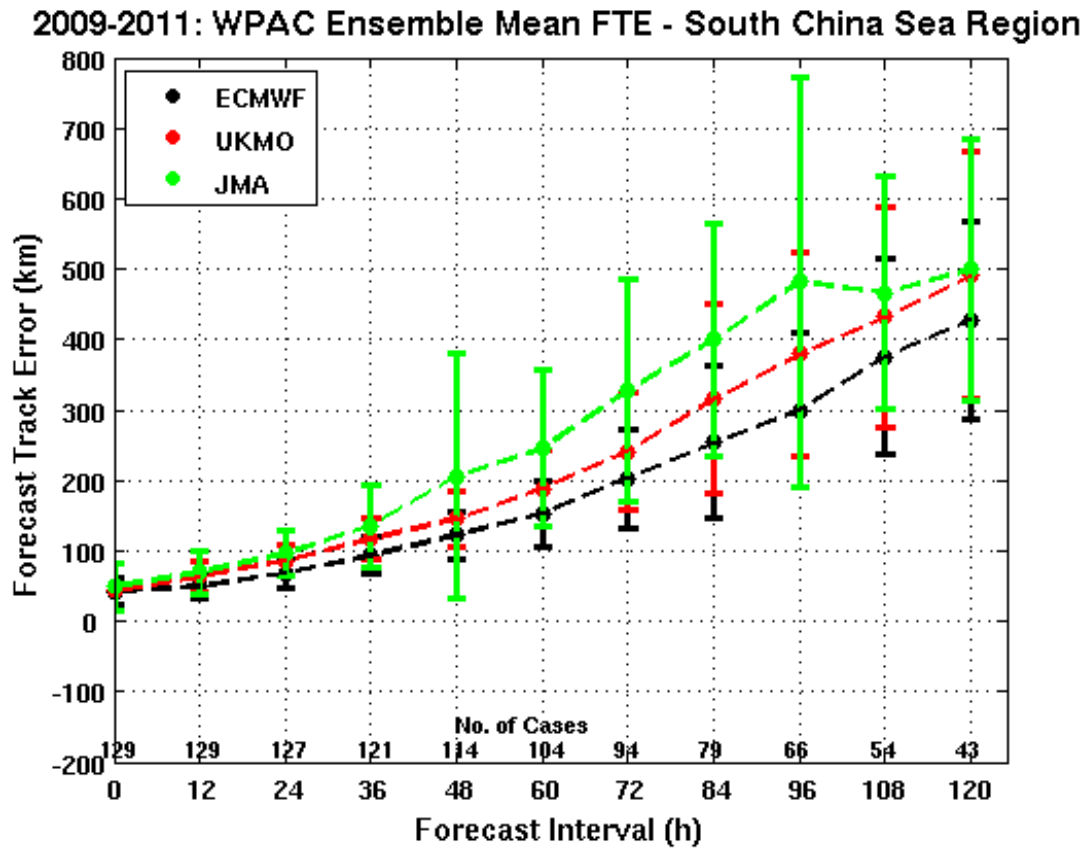


Figure 14. As in Figure 11, except for the SCSR region

The characteristics of ensemble-mean FTE over the Mid-Latitude Region (figure 15) are much different than the other regions, which is likely due to the small sample sizes and the highly variable flow patterns over the region. Small sample sizes may produce results based on one very large track error that can strongly influence the forecast-error or impact the magnitude of the standard deviation in the bar graph. The

ECMWF ensemble-mean FTE is smaller than that of UKMO and JMA (Figure 15). For the UKMO forecast intervals shorter than 84 h, the FTE over the MLR is larger than that over the entire WPAC (Figure 11). Beyond 84 h, the UKMO ensemble-mean FTE over the MLR is less than that over the WPAC. For ECMWF and JMA forecast intervals beyond 72 h, the ensemble-mean FTE over the MLR is larger than that over the entire WPAC.

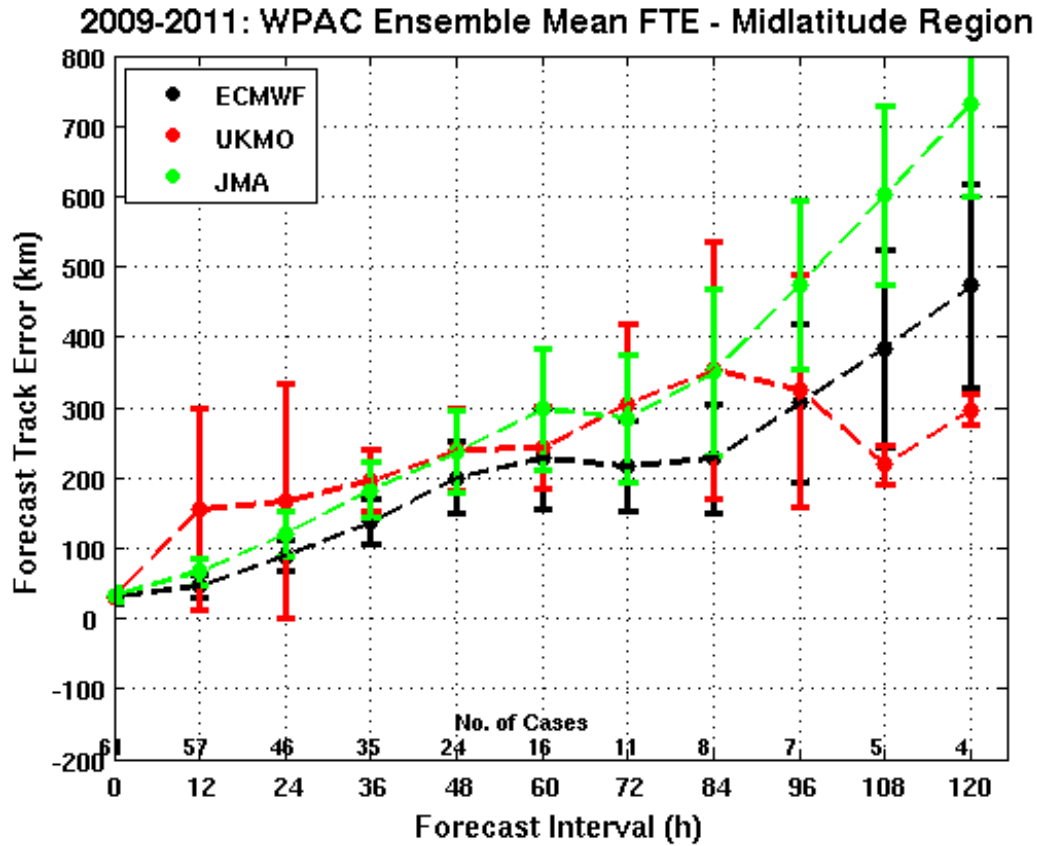


Figure 15. As in Figure 11, except for the MLR region

## 2. ATE

As defined in Chapter II, a positive (negative) Along-Track Error (ATE) indicates that the forecast position is ahead (behind) of the best-track position. The ATE reflects errors relative to advection of the TC along the correct path without regard for the flow orientation.

Over the WPAC, each individual ensemble-mean forecast has an average ATE between zero and 100 for each forecast interval. Therefore, the average ensemble forecast is ahead of the verifying best-track position (Figure 16). For each EPS, the standard deviations of the ensemble-mean ATE progressively increase as the forecast intervals increase.

A drastic increase occurs in standard deviation beyond 84 h. The increase in standard deviation is expected due to greater forecast uncertainty with increasing forecast interval; however, the magnitude of the standard deviation suggest high variability in long range TC forecasts as being ahead or behind the actual speed of movement. As will be clear when individual regions are examined, the standard deviation of medium-range ATE is likely due to large negative errors over the mid-latitude region (Figure 20). The ECMWF has the smallest ensemble-mean ATE and standard deviations about the mean ATE, which is consistent with the FTE results.

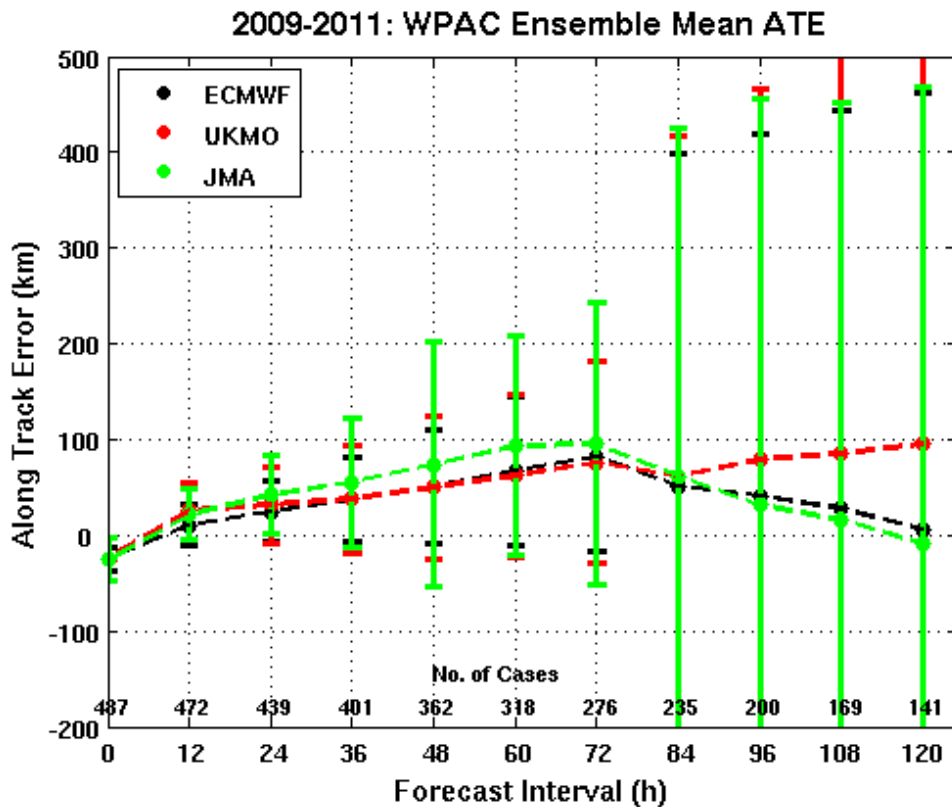


Figure 16. Average ATE for each of the EPS over the entire WPAC from 2009–2011. A plus and minus one standard deviation about the mean ATE is represented by the vertical line at each 12-h forecast from 0–120 h

Over the MTR (Figure 17), the ensemble-mean ATE for all EPS are also positive over all forecast intervals. There does not appear to be a significant timing bias in the ensemble-mean forecast over the MTR. However, the most significant difference between the MTR and the overall WPAC region is the lack of increased standard deviation over the MTR at larger forecast intervals.

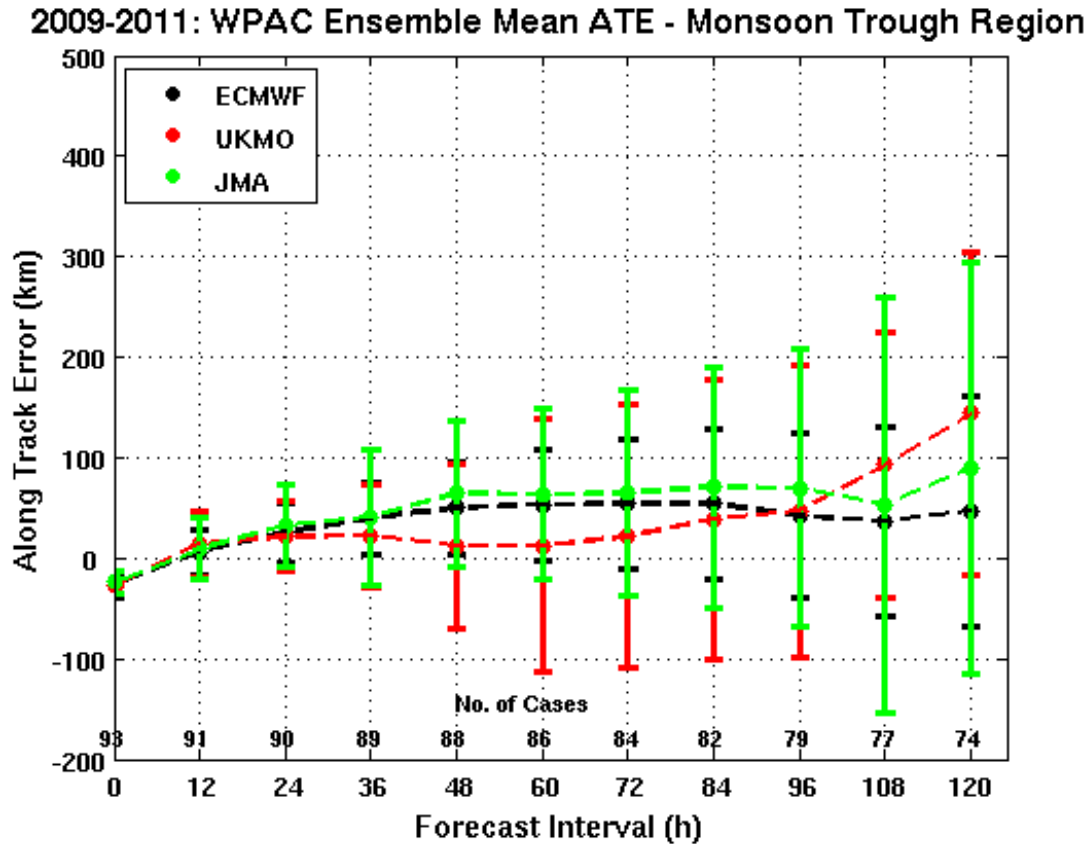


Figure 17. As in Figure 16, except for the MTR region

Over the RCR (Figure 18), the ensemble-mean ATE for all EPS are slightly more positive than over the MTR region. Also, the standard deviations about the mean ATE are larger, which reflects the variability associated with the TC recurvature. The ECMWF and JMA have negative ensemble-mean ATE over the medium-range forecasts. This indicates that ECMWF and JMA may be slower to forecast the recurvature of a TC at these forecast intervals.

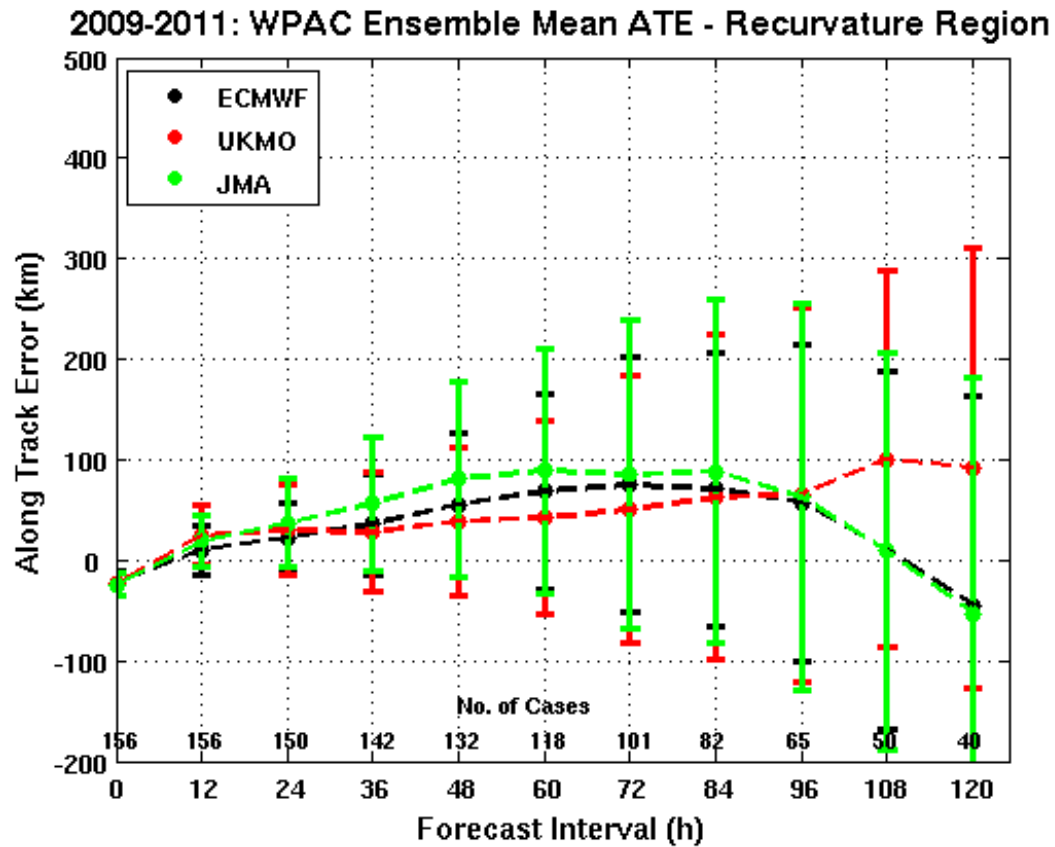


Figure 18. As in Figure 16, except for the RCR region

The largest positive ensemble-mean ATE for all EPS occurs over the SCSR (Figure 19). This may be due to increased variability in TC steering flow and erroneous forecasting of recurvature. The mean ATE forecast EPS are very similar such that no one EPS seems to be more accurate than any other.

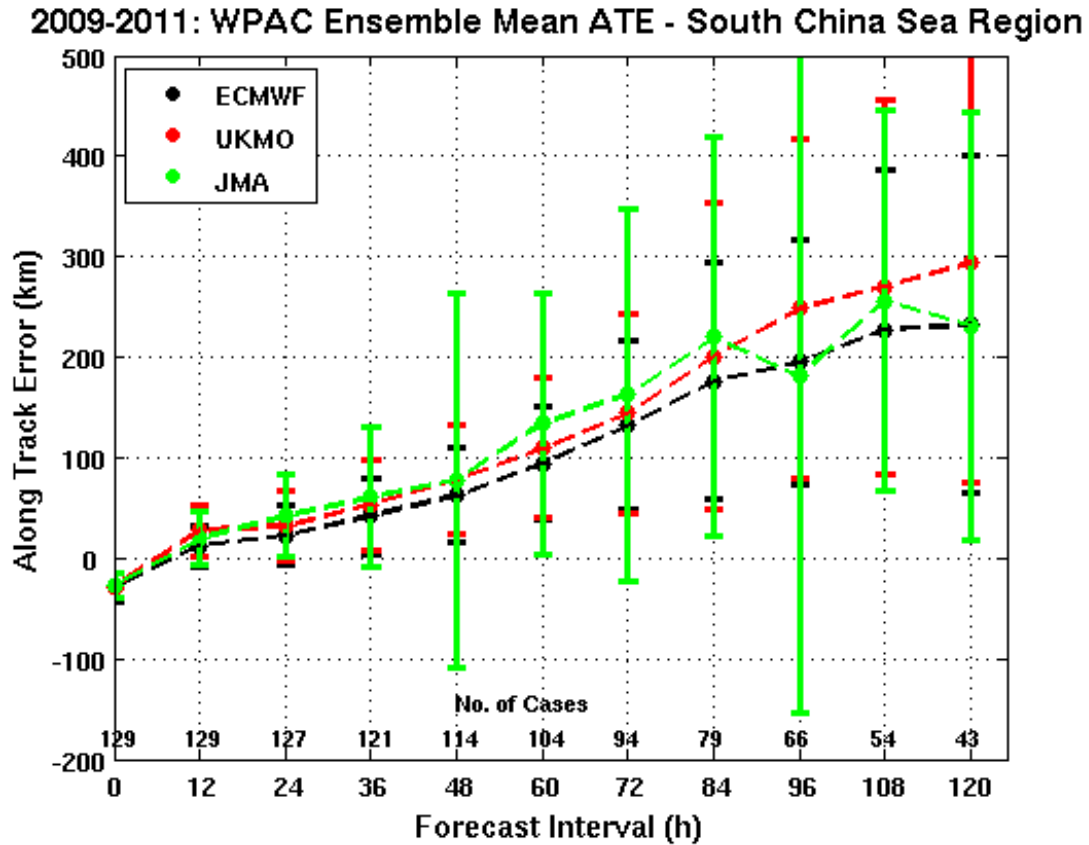


Figure 19. As in Figure 16, except for the SCSR region

Over the MLR (Figure 20), the ensemble-mean ATE for all EPS become very negative beyond 48 h. Therefore, the forecasts from all EPS are behind the actual TC tracks. This indicates that the EPS are not accurately forecasting the speed of TCs along the track as they enter the strong mid-latitude westerlies. Furthermore, the standard deviations about the mean ATE are small, which indicates that the speed bias in the ensemble-mean forecasts is quite systematic.

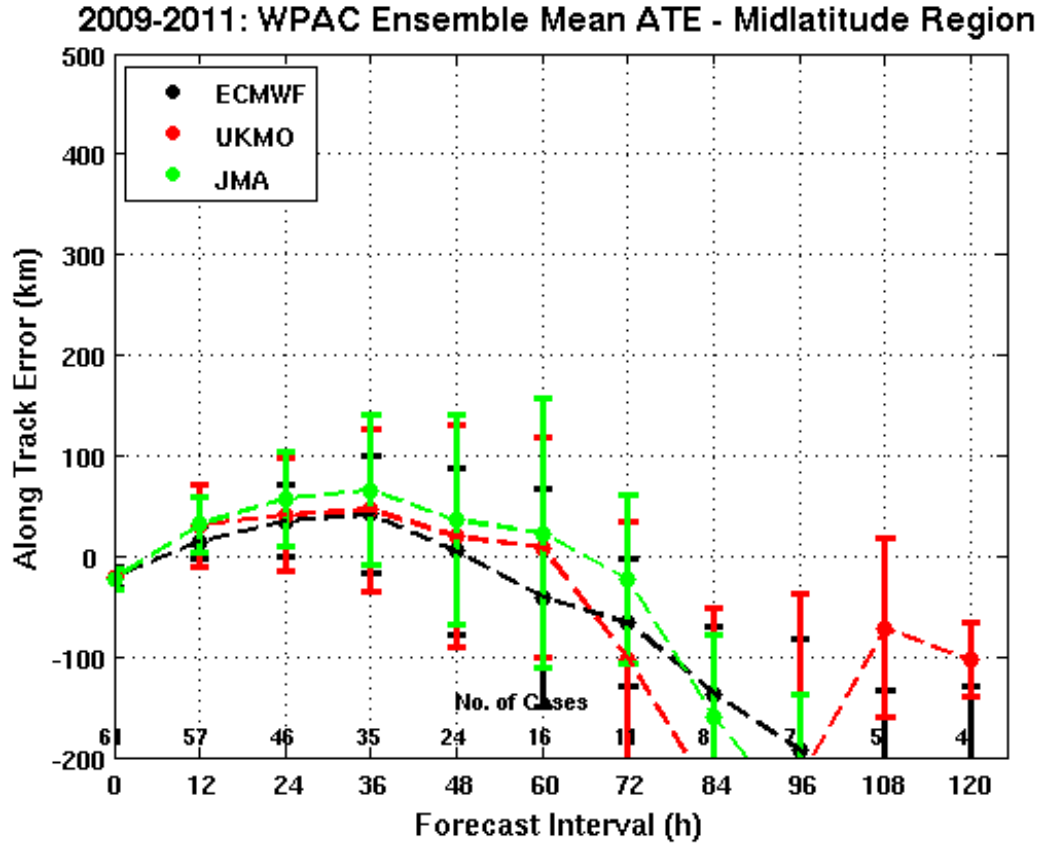


Figure 20. As in Figure 16, except for the MLR region

### 3. XTE

The Cross-Track Error (XTE) is the component of the FTE that indicates whether the forecast position is left (negative) or right (positive) of the best-track position. The XTE can indicate an error in predicting the orientation and speed in the motion of a TC relative to background steering flow.

For all EPS at forecast interval less than 72 h, the mean XTE is near zero, which indicates no clear bias in the ensemble-mean forecasts. The ECMWF and JMA ensemble-means have low average XTE across all forecasts intervals in the WPAC (Figure 21). The standard deviations about the mean XTE are small, which indicates consistency in predicting the orientation of the background flow. Beyond 72 h, the mean XTE becomes positive for each EPS, which indicates slight bias to the right of the best-track.

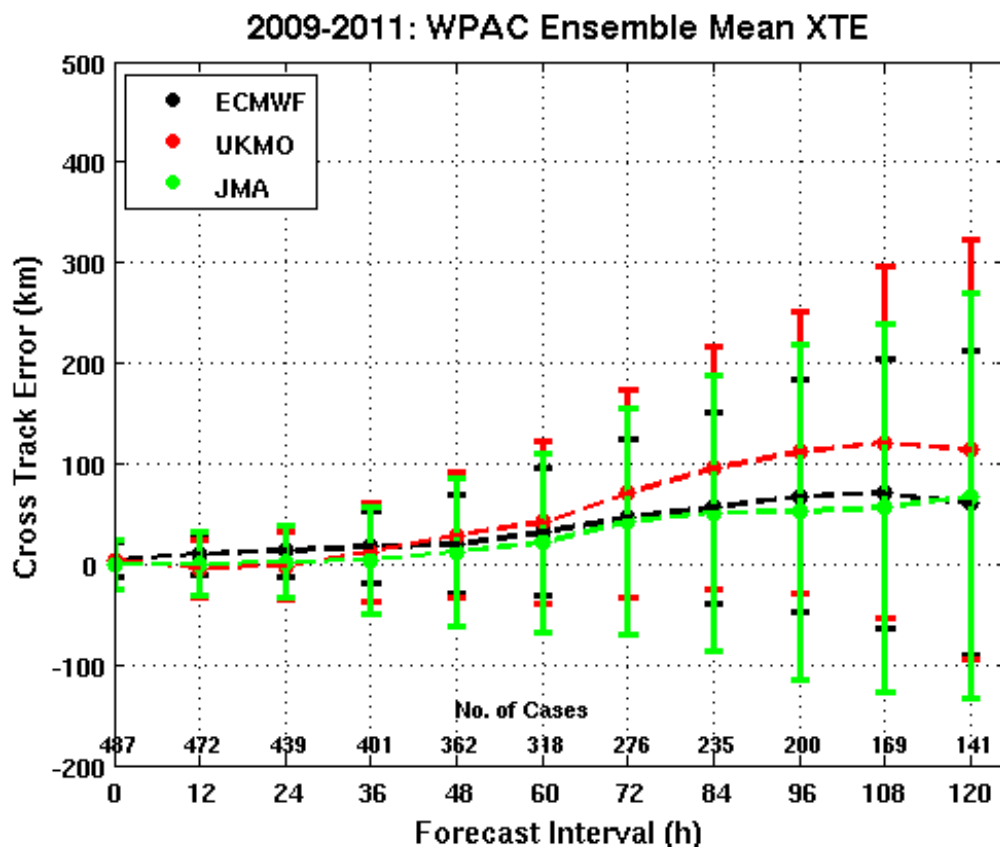


Figure 21. Average XTE for each of the EPS of the entire WPAC from 2009–2011. A plus and minus one standard deviation about the mean XTE is represented by a vertical line at each 12 h forecast from 0–120 h

Over the MTR (Figure 22), the ensemble-mean XTE characteristics are similar to the entire WPAC. At short forecast ranges, the mean XTE are near zero for all EPS. At longer forecast intervals, the mean XTE becomes positive, which indicates the same right bias as evident over the entire WPAC (Figure 21).



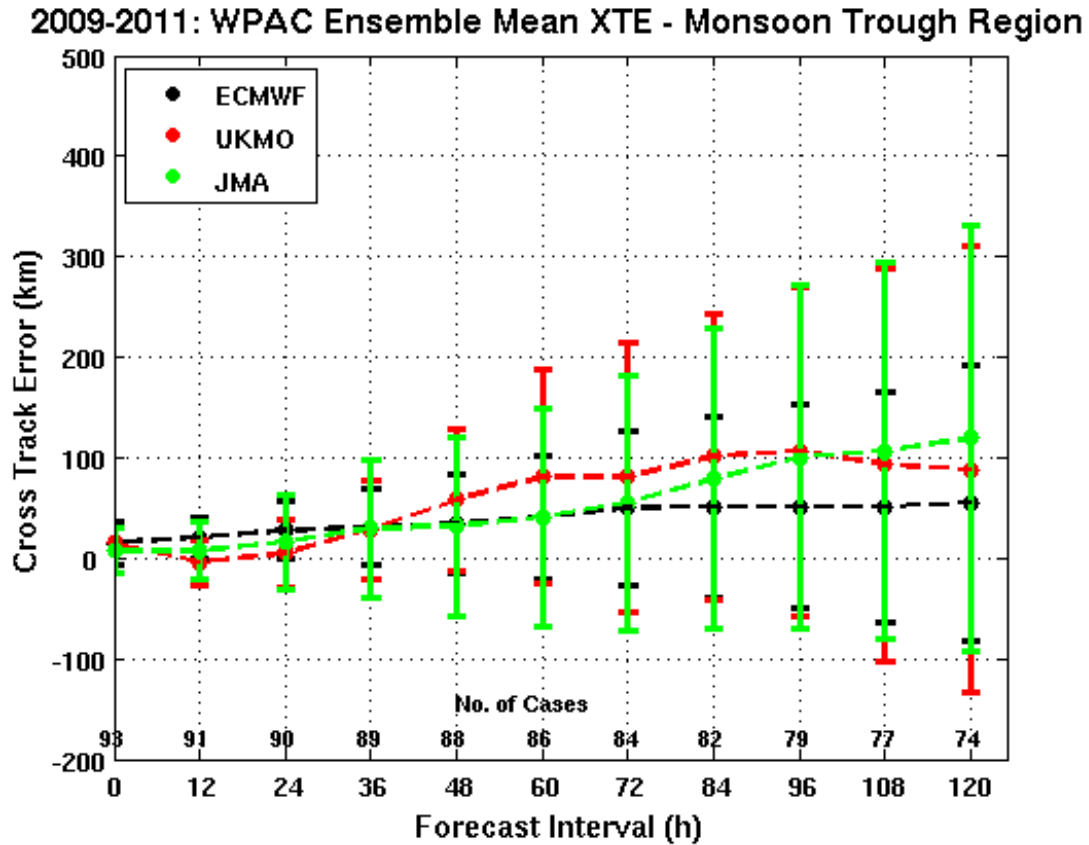


Figure 22. As in Figure 21, except for the MTR region.

Over the RCR (Figure 23), ensemble-mean XTEs become more positive than over the entire WPAC (Figure 21). The ensemble-mean XTE for the UKMO is much more positive than for JMA and the ECMWF, which are similar. This positive XTE bias indicates all three EPS consistently forecast tracks to the right of the best-track. This could be due to the EPS over-forecasting the influence of increasing mid-latitude westerlies as TCs tracks northward.

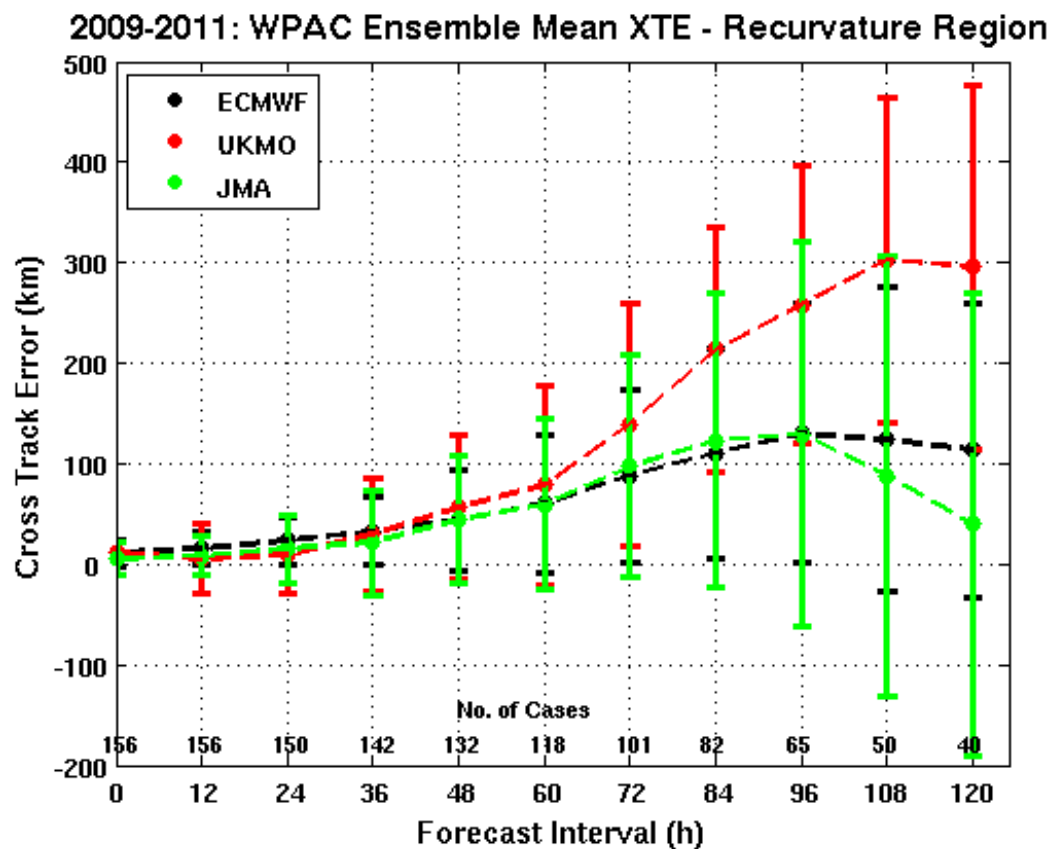


Figure 23. As in Figure 21, except for the RCR region

Over the SCSR sub-region (Figure 24) ensemble-mean XTE for all EPS are smaller than over any other regions. On average, there is no defined bias in the ensemble-mean forecast tracks.

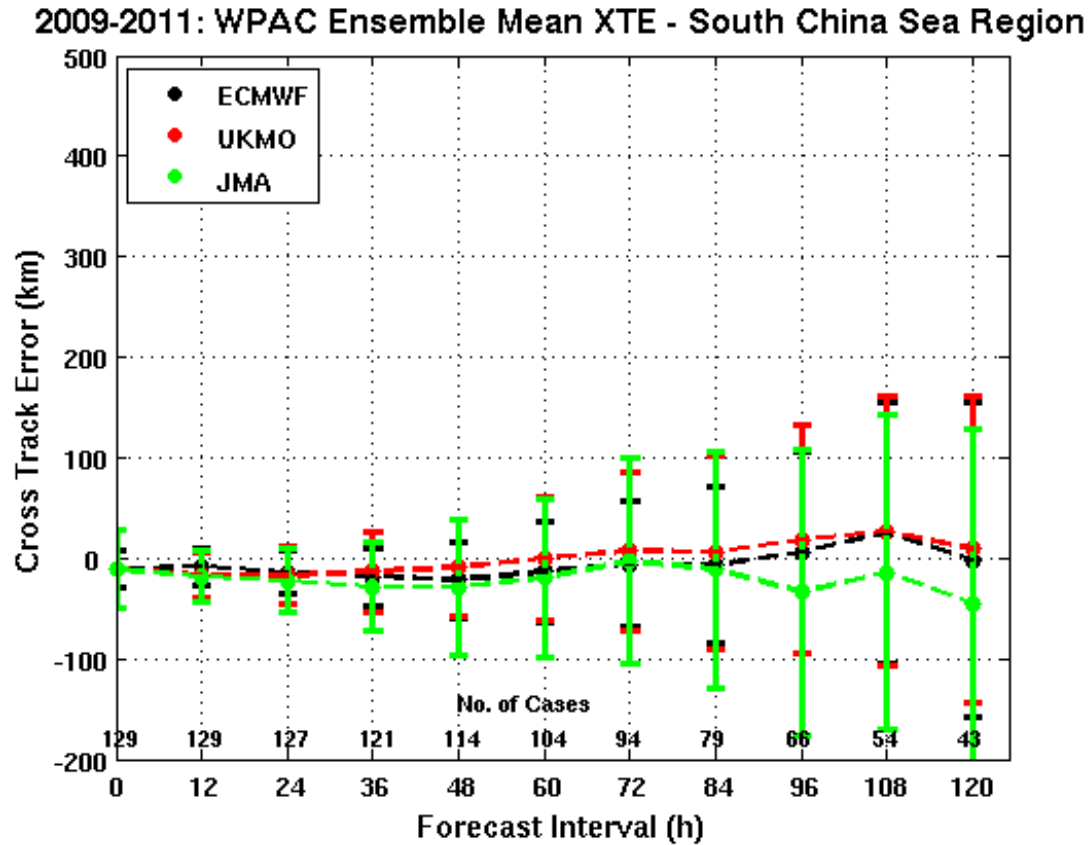


Figure 24. As in Figure 21, except for the SCSR region

Consistent with the ATE over the MLR region (Figure 25), the ensemble-mean XTE for all EPS becomes very negative at larger forecast intervals. Therefore, all EPS have large bias. This is consistent with the negative ATE errors. As a TC turns eastward under the influence of increased mid-latitude westerlies, the storm is forecast to move too slow, which would place it to the left (west) of the best-track.

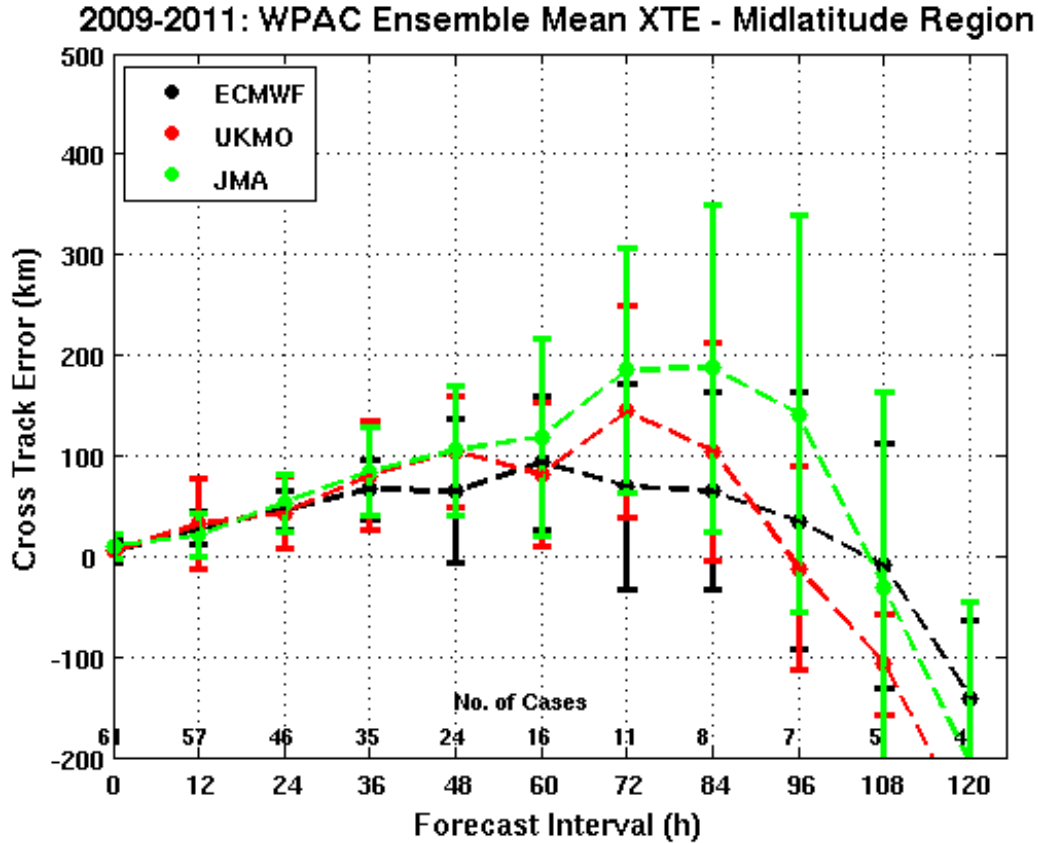


Figure 25. As in Figure 21, except for the MLR region

### C. PROBABILITY WITHIN SPREAD

The Probability within Spread (PWS) is a calculation that defines the proportion of time that the best-track is contained within the spread of the individual ensembles. Spread is defined as being one standard deviation to the right or left of the ensemble-mean track as defined by the individual ensemble members. An EPS that has a high (low) PWS for a particular forecast interval indicates that the individual ensemble forecast track members do (not) have enough spread to reflect the track uncertainty. For the purposes of this study, a PWS value of 0.68 for  $1\sigma$  was utilized to indicate statistical consistency.

Throughout the intermediate forecast intervals, each EPS has a PWS that is between 0.4 and 0.5 over the WPAC (Figure 26). The PWS is small at the initial time of

the forecast due to small perturbations in the initial conditions. The initial condition uncertainty leads to large spread at very short forecast intervals. The PWS increases rapidly at 12 h in response to the initial condition uncertainties.

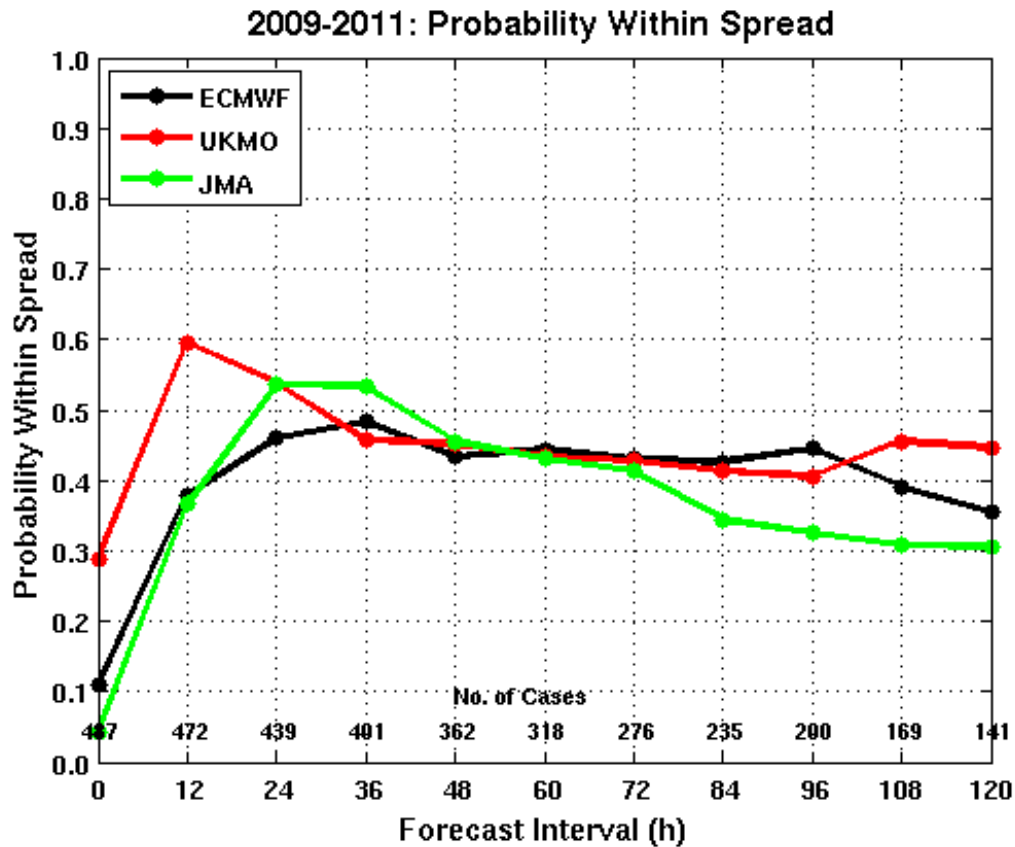


Figure 26. The PWS for each EPS over the entire WPAC for the 2009–2011 seasons

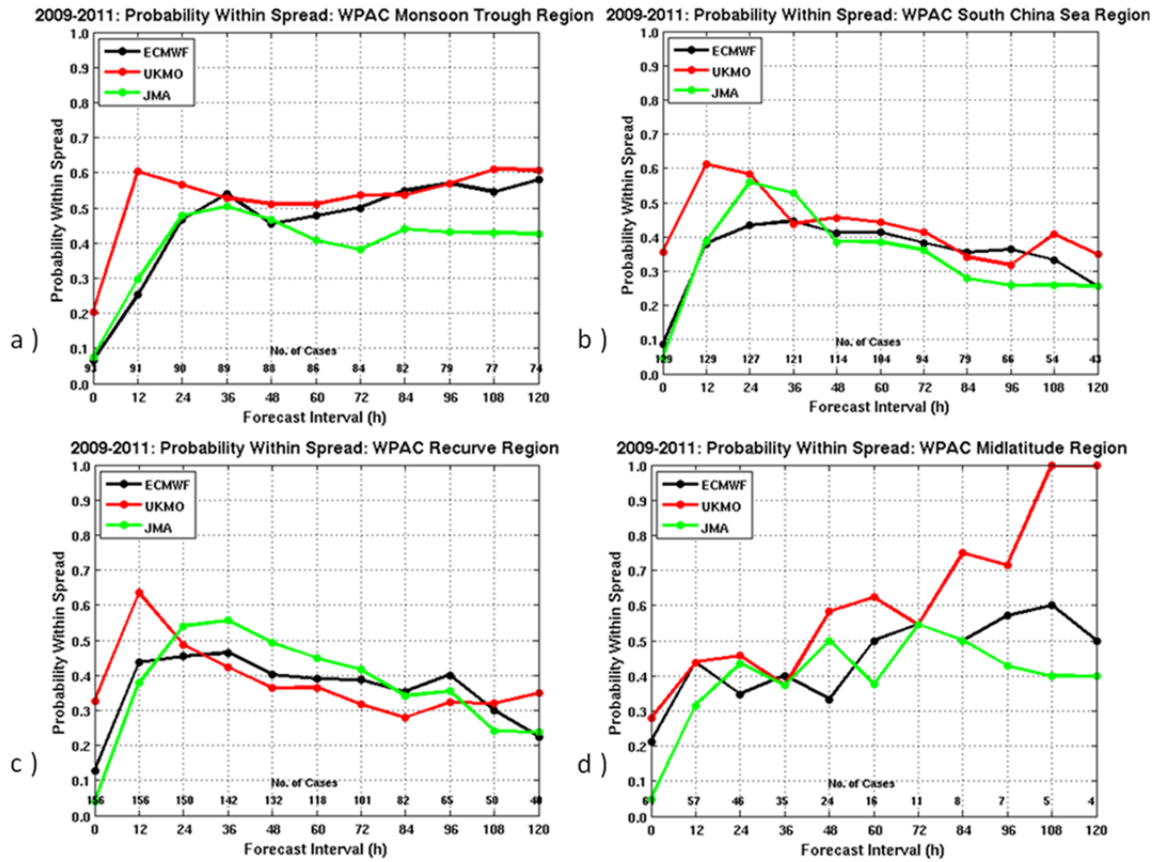


Figure 27. The PWS for each EPS over each sub-region of the WPAC for the 2009–2011 seasons

Over the MTR (Figure 27a), the PWS was larger than the basin average for all EPS. However, over the SCSR (Figure 27b) and the RCR (Figure 27c), the PWS was smaller, which indicates that spread was too small with respect to potential recurvature tracks. Over the MLR (Figure 27d) the PWS is larger than the RCR and SCSR. This is due to two reasons. There are relatively few numbers of cases or homogeneous forecasts members over the MLR. Also, the spread over the MLR is generally much larger than the other regions due to the increased variability in mid-latitude westerlies. Therefore, the PWS is high, but the resolution is likely to be low.

#### D. ELLIPSE RELIABILITY

In this thesis, reliability is defined as the percentage of time that the best-track analysis position is contained within the EPS ellipse at a particular forecast time. As discussed in Chapter III, the EPS ellipse is defined as enclosing 68% of the ensemble

forecast track members, based on a Chi-squared scaling. The reliability percentages of the EPS which are higher (lower) than 68% will be determined to be under (over) dispersive.

In Figures 28 and 29, the reliabilities of individual EPS ellipses are shown by line graphs for each forecast interval. The blue line symbolizes the expected reliability of the ellipses at 68%. The values above the forecast interval are the number of EPS forecasts that were included.

The reliability, or ellipse hit rate, for all of EPS is below 68% for the WPAC (Figure 28) and all sub-regions (Figure 29). The ECMWF has the highest average reliability of the three EPS over the entire WPAC (Figure 28). The distribution of reliability over the four sub-regions is similar to that of PWS. That is, the reliability over the MTR and SCSR is higher than the reliability over the entire WPAC. Ellipse reliability over the RCR is smaller than over the entire WPAC. Although, the PWC over the MTR is larger, the reliability is the smallest of all regions.

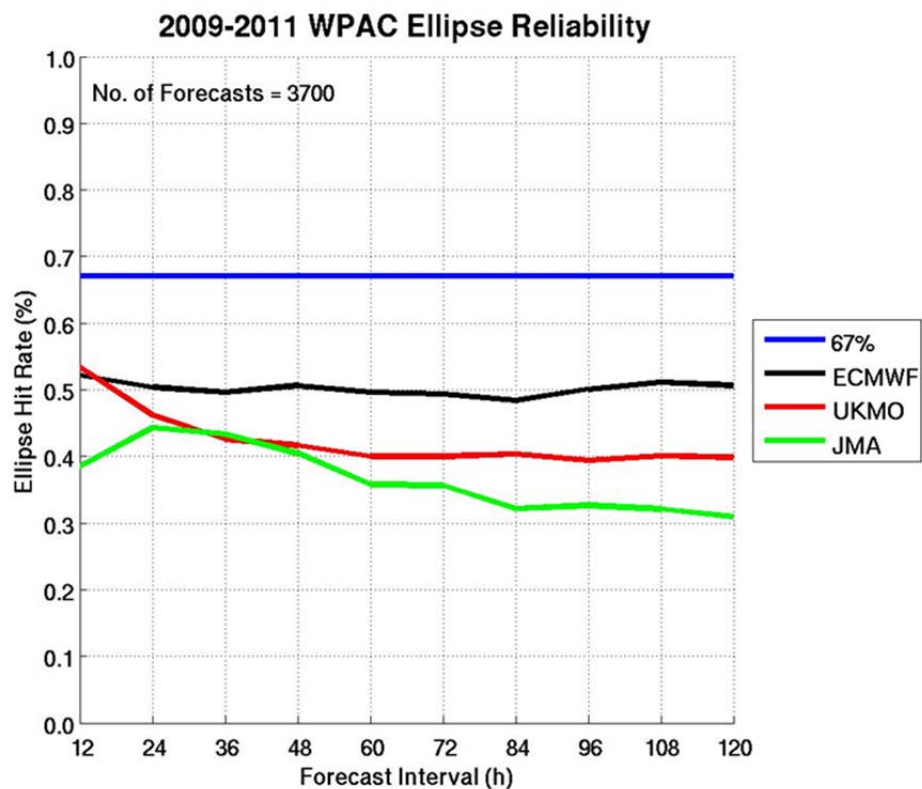


Figure 28. The ellipse reliability for each EPS for the entire WPAC

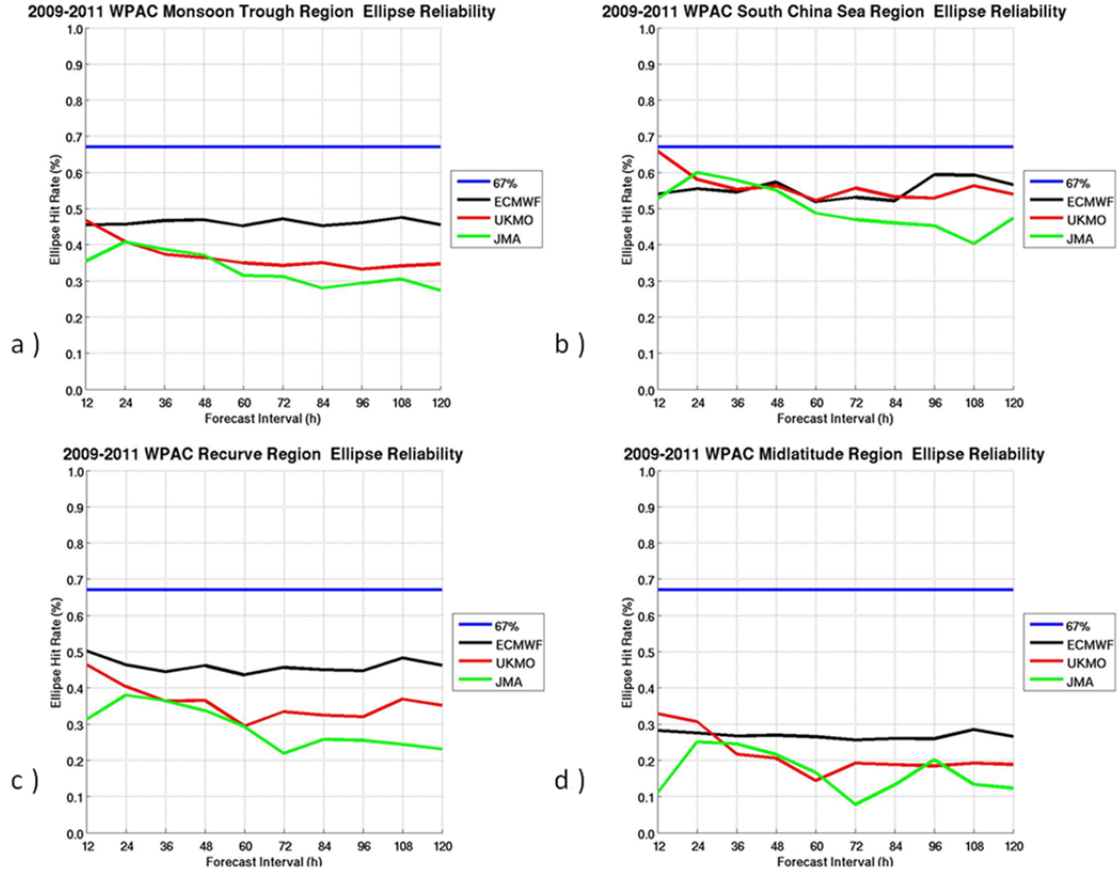


Figure 29. The ellipse reliability for each EPS over the each sub-region of the WPAC for the 2009–2011 seasons as defined in Figure 28

### E. MEAN AREA DIFFERENCE (MAD)

Because the ensemble-mean forecasts for the ECMWF generally exhibit the smallest FTE, ATE, and XTE, the MAD calculation is based on the ellipse area of the ECMWF. The MAD value measures the sizes of the UKMO or JMA ellipses relative to the ECMWF ellipse. A positive (negative) value indicates that the UKMO or JMA ellipse is smaller (larger) in size than the ECMWF ellipse. This would indicate that the UKMO or JMA predicts a reduced (higher) level of uncertainty. The values above the forecast interval show the number of EPS forecasts included.

Over all forecast intervals the MAD for the UKMO and JMA with respect to the ECMWF is negative (Figure 30) for the entire WPAC. At forecast intervals beyond 60 h, the MAD values are consistently near -1.00 for the UKMO and near -2.5 for JMA. Therefore, the JMA ellipse sizes are the largest of all three EPS investigated in this thesis.



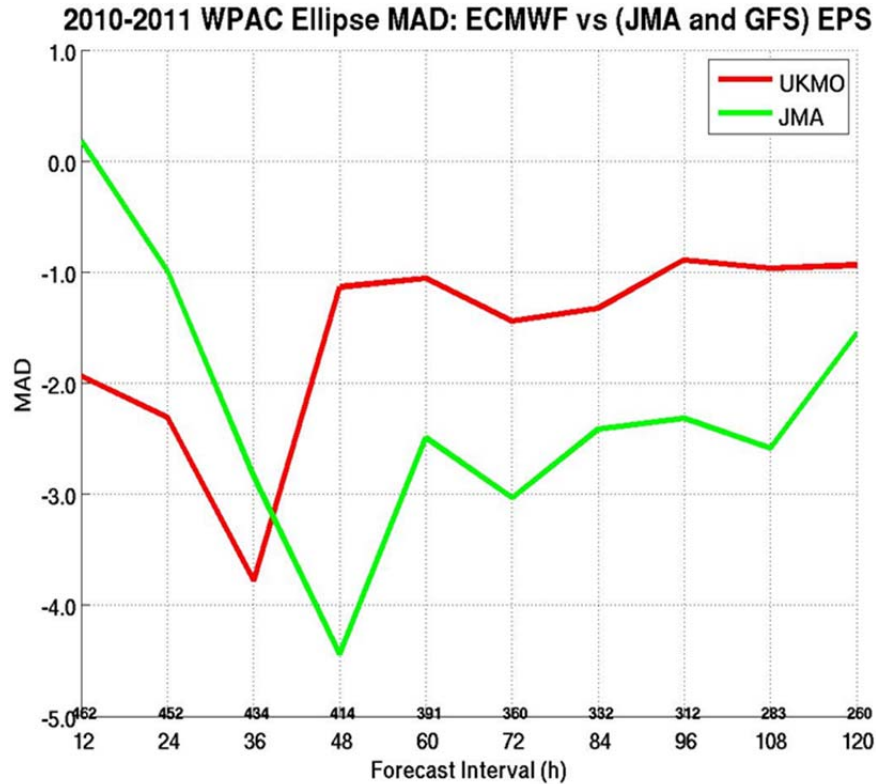


Figure 30. The Mean Area Difference (MAD) of UKMO or JMA EPS and compared to ECMWF across the WPAC. Positive values indicate the EPS ellipses are smaller than the ECMWF circle for each forecast interval. The values above the forecast interval are the number of EPS forecasts included

## F. SUMMARY

The objective of examining TC track errors from individual ensemble models is to enhance and analyze the attributes and utility of forecast reliability and resolution with respect to the ensemble spread as an aid to TC forecasting. Overall, errors were smallest for ECMWF. Additionally, the standard deviations about the mean error were smallest for ECMWF. ATE over the entire WPAC indicates average ensemble forecasts are ahead of the verifying best-track position. The average XTE over the entire WPAC for forecasts beyond 72 h indicates a slight bias to the right of the best track.

In the MTR there is a slight right bias. However, this contains less error variability when compared to other regions in which steering flows may be more complex steering.

In the RCR, it is evident that the variable steering flow and forecasted TC recurvature is influenced by mid-latitude flow patterns. The EPS tend to over-forecast the influence of increasing mid-latitude westerlies as TCs track northward. Of note, ECMWF and JMA may be slower to forecast the recurvature of a TC at these forecast intervals.

In the SCSR, the increased FTE and larger standard deviations for the JMA ensemble mean could indicate poor forecasts of flow patterns over the SCSR leading to early recurvature of the TCs in the mid-range forecasts. The UKMO model forecasts are the most accurate over the SCSR based on FTE analysis. However, there is no defined ATE or XTE bias for any EPS.

In the MLR, the characteristics of ensemble-mean FTE are much different than the other regions, which is likely due to the small sample sizes and the highly variable flow patterns. The ECMWF ensemble-mean FTE is smaller than that of UKMO and JMA. The ATE for all EPS indicates that forecast positions are behind the actual TC tracks. This indicates that the EPS are not accurately forecasting the speed of TCs along the track as they enter the strong mid-latitude westerlies. Furthermore, the standard deviations about the mean ATE are small, which indicates an along track bias, or bias in translation of speed. The XTE is consistent with the ATE in that all EPS have large bias to the right of the track, which is consistent with negative ATE.

The PWS is small and relatively similar for each EPS throughout the WPAC, which indicates that the probability of the observed track being within the ensemble spread was small. There is no clear distinction as to which EPS had the highest PWS.

The ellipse reliability for all of EPS is below 68% for the WPAC and all sub-regions, which indicates under-dispersion. The ECMWF has the highest average reliability of the three EPS over the entire WPAC. The distribution of reliability over the four sub-regions is similar to that of PWS. That is, the reliability over the MTR and SCSR is higher than the reliability over the entire WPAC.

The negative MAD values for the UKMO and JMA indicate that these EPS result in ellipses that are larger in size than the ECMWF ellipse. Thus, the UKMO or JMA may

overestimate uncertainty. The increased size of the UKMO and JMA ellipses, as compared to the ECMWF, also indicate the decrease of sea maneuverability when relying on UKMO and JMA to establish visual aids that convey TC track areas of uncertainty.

In summary, the statistical analyses which proved most informative were the basin and sub-regional track-errors and the MAD. The track-errors conveyed amplifying information about which individual EPS performed best over the basin and in specific-sub-regions. The UKMO ensemble showed merit in the forecasting-track error in the SCSR, while the JMA ensemble did not clearly distinguish itself in any of the sub-regions. The PWS and MAD, along with the majority of the track-error analysis, conclude that the ECMWF ensemble consistently outperforms the UKMO and JMA ensembles over a majority of the WPAC and the sub-regions, when comparing the attributes of reliability, resolution, and sharpness of forecasts.

#### **G. CASE STUDY TYPHOONS NANMADOL AND SONGDA**

Typhoon Nanmadol was chosen in this case study because of the high level of uncertainty surrounding its forecast track and the potential to impact Okinawa and mainland Japan (Figure 31). Typhoon Songda was chosen to contrast with Typhoon Nanmadol because it has a lower level of uncertainty (Figure 32) and forecasts were generally along the best-track. The initial JTWC track forecasts for both Songda and Nanmadol were both very close to Okinawa.

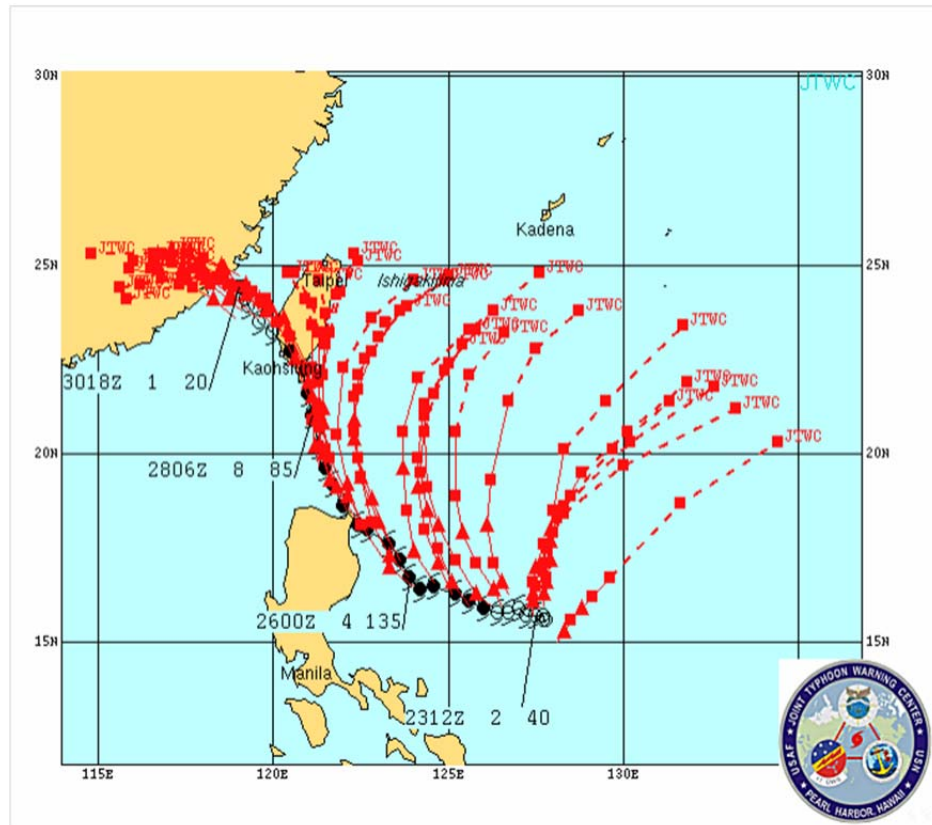


Figure 31. All JTWC track forecasts for Typhoon Nanmadol. The red dashed lines correspond to each JTWC forecast track. The best-track positions are in black (From: JTWC 2012a)

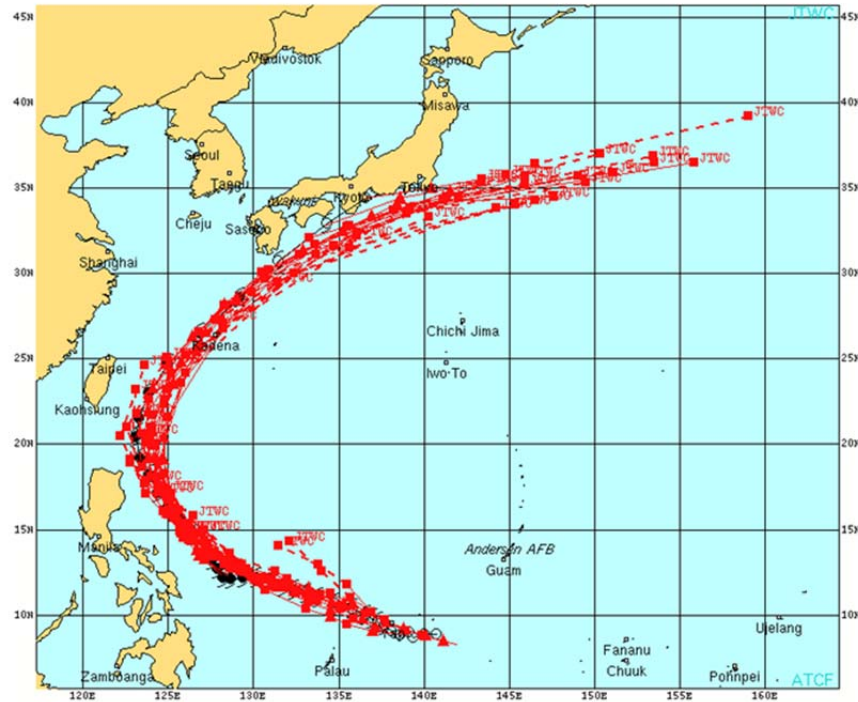


Figure 32. All JTWC track forecasts for Typhoon Songda. The red dashed lines correspond to each JTWC forecast track. The best-track positions are in black (From: JTWC 2012a)

## 1. Typhoon Songda

On 22 May 2011, Super Typhoon Songda formed east of Palau (Figure 32) and rapidly intensified to a peak of 140 kt by 27 May as it recurved east of the Philippines and Taiwan (JTWC 2012a). By 28 May, Songda subsequently weakened to 80 kt under the influence of increasing vertical wind shear as it passed approximately 40 nautical miles to the north-northwest of Okinawa (JTWC 2012a). On 29 May, the cyclone passed along the southern coast of Honshu before completing extra-tropical transition and accelerating eastward into the central North Pacific as a baroclinic low pressure system (JTWC 2012a).

### a. Typhoon Songda 1200 UTC 23 May 2011

The Typhoon Songda case study begins 1200 UTC 23 May 2011 (Figure 31). The ECMWF was the only EPS that had significantly smaller along- and cross-track components, and a corresponding small spread of ensemble members. This is consistent

with the summaries of XTE and ATE in the MTR (Figure 17 and 22) and RCR (Figure 18 and 23). The ensemble members of UKMO also had very little spread in the individual forecast members, but the ellipses were larger in both along- and cross-track components in the short range when compared to ECMWF. The UKMO forecast spread increased over the mid- to long-range forecasts near the location of forecast recurvature. The JMA ensemble spread was the greater over all the forecast intervals. The ellipse orientation of the JMA EPS in the short-range forecast has a large cross-track component, and in the long-range they show a long along-track component. This indicates that JMA predicted the greatest uncertainty as to when Songda will begin to recurve.

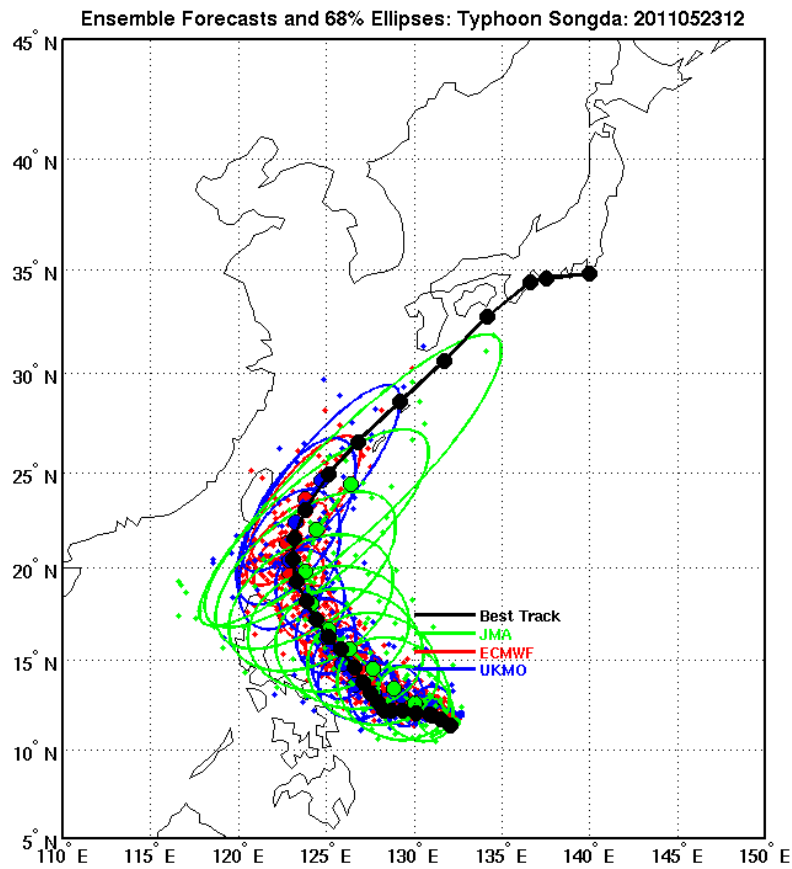


Figure 33. The TC forecast track ellipses of each EPS for Typhoon Songda 1200 UTC 23 May 2011. Each ellipse signifies a 12 h forecast interval and is colored to match the individual EPS as defined in the legend at the top right. The large dot inside each ellipse is the corresponding ensemble-mean forecast position. The best-track positions are in black

***b. Typhoon Songda 1200 UTC 24 May 2011***

For the forecast initiated at 1200 UTC 24 May 2011 (Figure 33), the ellipse formations before recurvature were smaller than the previous forecast sequence. However, the largest area of uncertainty for the forecast period was post recurvature when the ellipses became elongated along the forecast track. Therefore, the primary uncertainty is associated with the timing of the storm as it moved into the mid-latitude region.

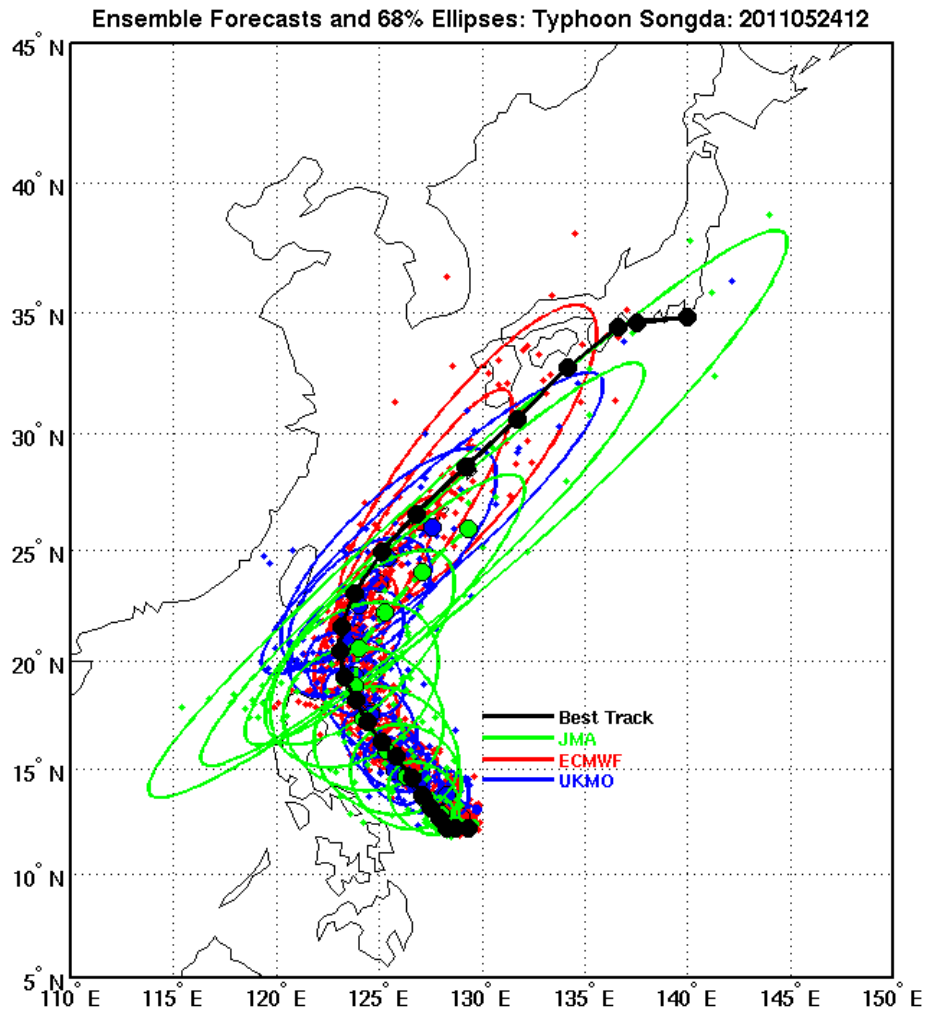


Figure 34. The TC forecast track ellipses of each EPS for Typhoon Songda 1200 UTC 24 May 2011, as defined in Figure 33

*c. Typhoon Songda 1200 UTC 25 May 2011*

By 1200 UTC 25 May 2011 (Figure 33), it was clear that the primary uncertainty in the forecast of Songda was in the speed of motion along the track. All three EPS exhibited less uncertainty in the short-range forecasts.

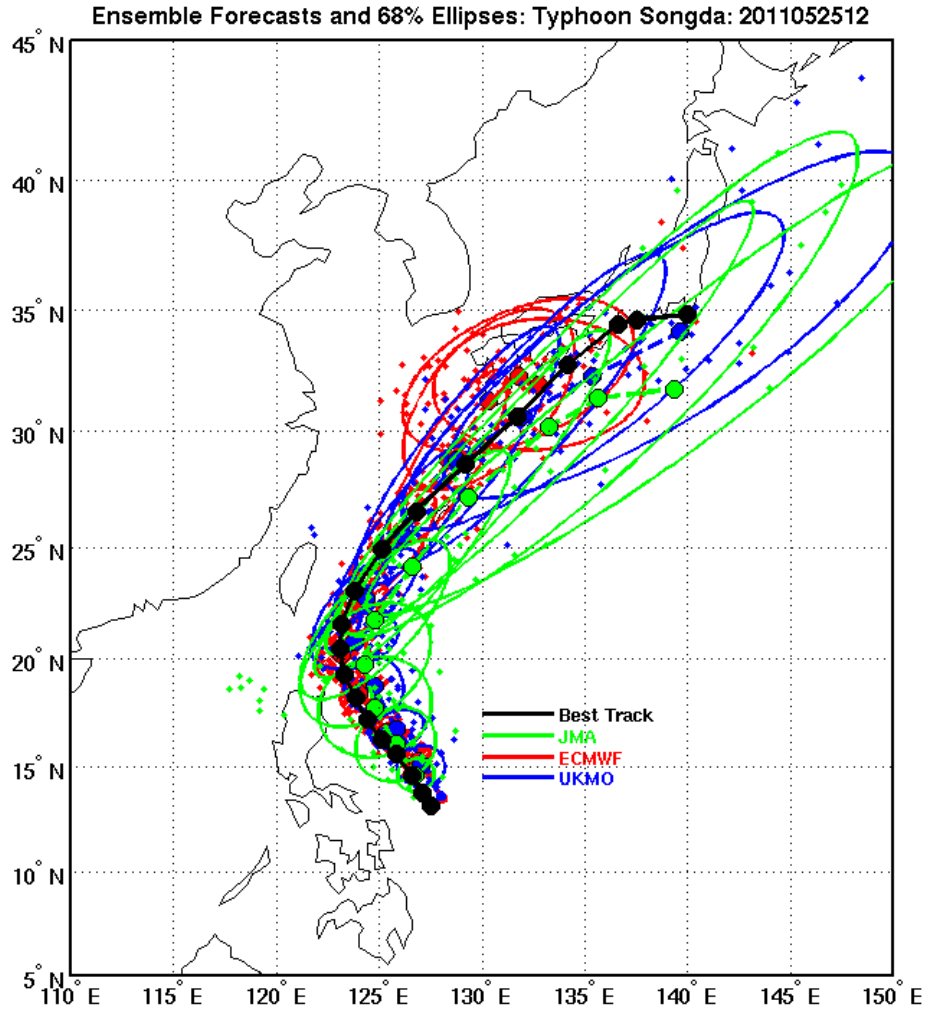


Figure 35. The TC forecast track ellipses of each EPS for Typhoon Songda 1200 UTC 25 May 2011, as defined in Figure 33

## 2. Typhoon Nanmadol

On 23 August 2011, Super Typhoon Nanmadol (Figure 31) formed within the monsoon trough east of the Philippines and began tracking west-northwestward toward



Luzon in a complex steering environment dominated by a subtropical ridge to the north and east (JTCWC 2012a). On 25 August, the storm took a poleward turn around the steering ridge and rapidly intensified to reach super typhoon status by 26 August under the favorable environmental influences of low vertical wind shear, excellent dual-channel outflow enhanced by a Tropical Upper Tropospheric Trough (TUTT) cell to the northeast, and passage over a region of high ocean heat content (JTCWC 2012a). On 26 August, Typhoon Nanmadol clipped the northeast tip of Luzon and then moved across the southern coast of Taiwan before dissipating in the Taiwan Strait (Figure 31), just prior to making landfall in China's Fujian Province (JTCWC 2012a).

*a. Typhoon Nanmadol 1200 UTC 23 August 2011*

Predicted across-track uncertainty for Typhoon Nanmadol was high for all three EPS (Figure 36). Overall, the ellipses were elongated in a zonal, or across track fashion. The ECMWF and JMA ellipses did not contain the best-track positions. However, the UKMO had the largest ellipses. Interestingly, the UKMO ensemble-mean forecast track was closest to that of the actual TC track, but had the largest spread. This is consistent with the XTE summary for the SCSR (Figure 24) where UKMO had the smallest average XTE.

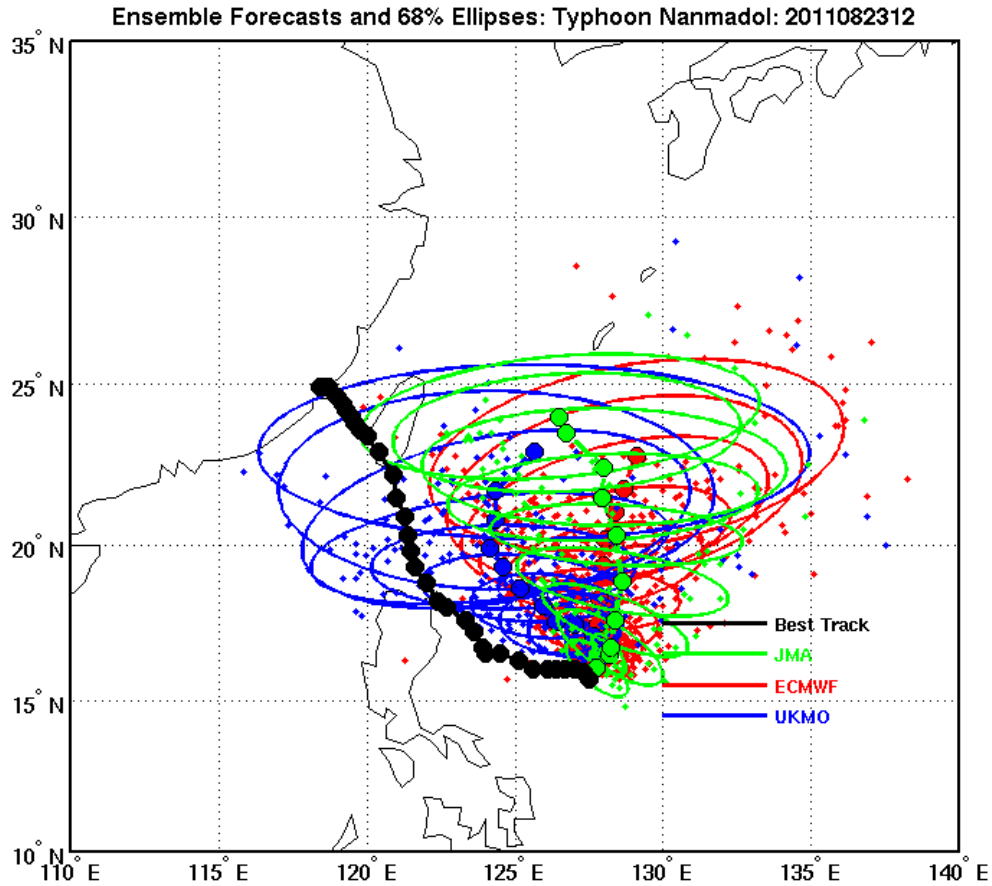


Figure 36. The TC forecast track ellipses of each EPS for Typhoon Nanmadol 1200 UTC 23 August 2011. Each ellipse signifies a 12 h forecast interval and is colored to match the individual EPS as defined in the legend at the top right. The large dot inside each ellipse is the corresponding ensemble-mean forecast position. The best-track positions are in black

***b. Typhoon Nanmadol 1200 UTC 24 August 2011***

At 1200 UTC 24 August 2011 (Figure 37), Nanmadol progressed west towards Luzon, and the ensemble forecasts begin to shift more towards the best-track

positions. The three EPS continue to exhibit uncertainty in across track error. At longer forecast intervals the JMA ellipses are becoming parallel to the best-track, which indicates uncertainty in the speed of advance.

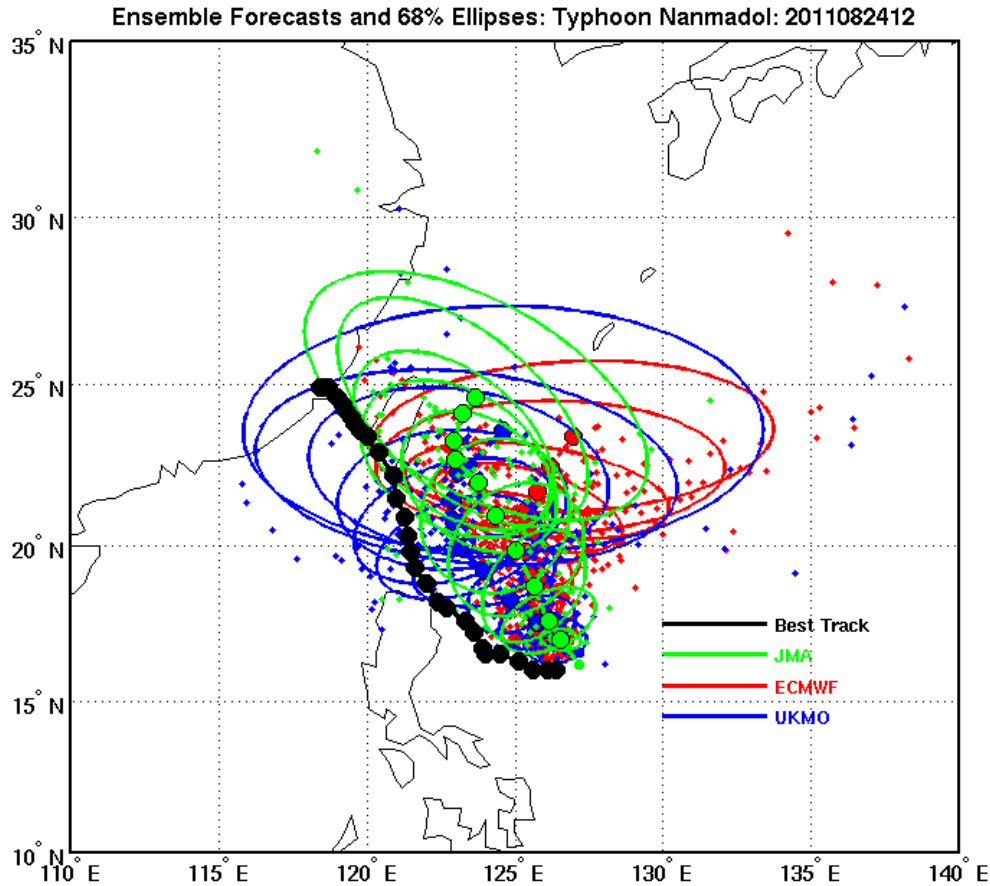


Figure 37. The TC forecast track ellipses of each EPS for Typhoon Nanmadol 1200 UTC 24 August 2011, as defined in Figure 36

*c. Typhoon Nanmadol 1200 UTC 25 August 2011*

At 1200 UTC 25 August 2011 (Figure 38), Nanmadol had progressed westward towards Taiwan. The ensemble-mean track of all three EPS continues to predict a turn towards Okinawa. At longer forecasts intervals, the UKMO had the largest

ellipses, but the ensemble mean forecast track was the closest to the best-track positions in the short- to mid- range. The UKMO ellipses orientation in the long range were turned more parallel to the best-track positions indicating it had difficulty accurately forecasting the TC's speed of advance. JMA also showed uncertainty with larger ellipses oriented across track in the short-range, and along-track in the long-range speed of advance. The ellipses of the ECMWF EPS were the smallest in the short- to mid-range, but continue to be shifted to a forecast track that is too far east.

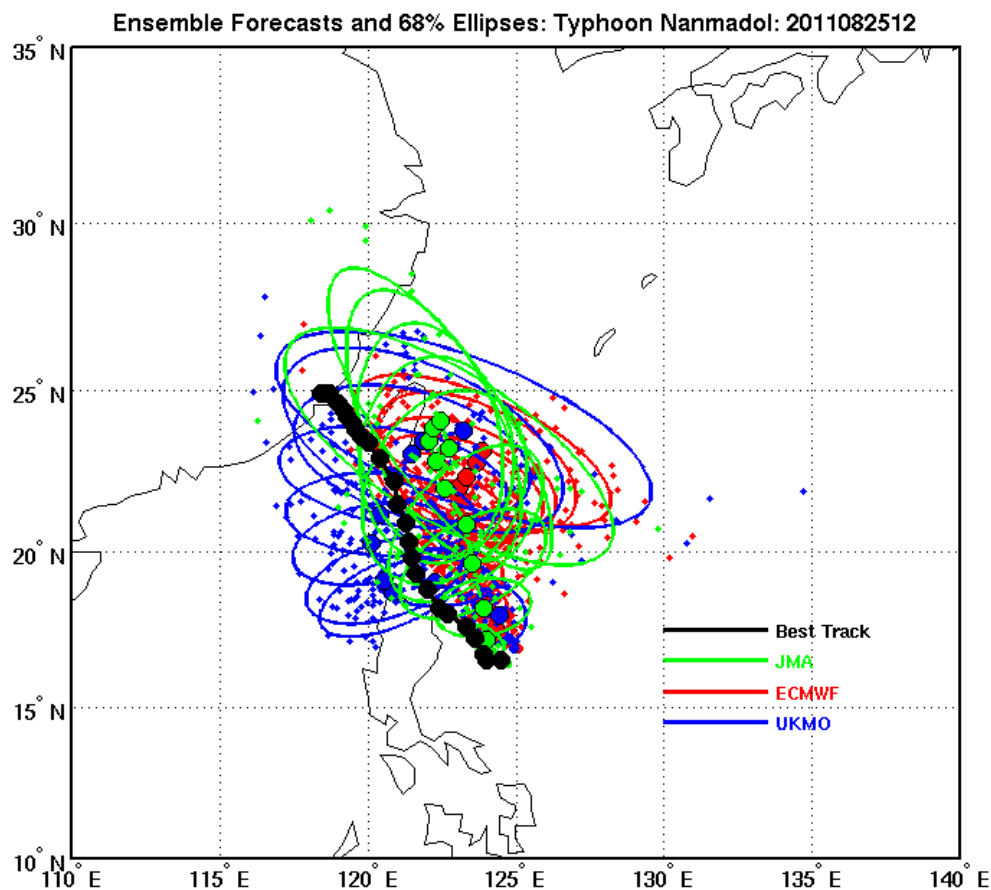


Figure 38. The TC forecast track ellipses of each EPS for Typhoon Nanmadol 1200 UTC 25 August 2011, as defined in Figure 36

### **3. Summary**

Typhoons Nanmadol and Songda were selected as case studies because of their similar initial forecast tracks, contrasting levels of uncertainty, and impending sortie and COR decisions that would have to be recommended by Fleet Oceanographers and made by regional, installation, and unit commanders. The decision to sortie or not sortie assets is a convincing justification of why probability products are relevant in portraying forecast confidence to Fleet Oceanographers and Commanders.

In the case of Typhoon Nanmadol, the EPS ellipses were elongated normal to the forecast and best-tracks, which suggests the potential for across-track errors. Operational forecasters who are able to view figures that show ellipse orientation similar to the case of Typhoon Nanmadol should be able to comprehend that there is spatial uncertainty in the models forecasted track direction. The ellipses in the case of Songda suggested that potential along-track errors, which are parallel to the forecasted and best-tracks. This case should convey to a forecaster that there is temporal uncertainty in the speed at which the storm is forecasted to move along-track.

Consensus-based forecasts and individual numerical models used to produce visual decision aids did not demonstrate positive skill in forecasting the cross-track variability that was observed in the case of Typhoon Nanmadol. In the case of Typhoon Songda, it did not demonstrate positive skill in forecasting along-track in the RCR and MLR. The use of EPS ellipses could have enhanced the representation of uncertainty of the forecasts for both cases.

Compared to the FTE analysis results in the SCSR, the UKMO ellipses for Typhoon Nanmadol were the most reliable in capturing the best-track positions. However, UKMO ellipses did have lower resolution in the mid- to long-range forecast intervals indicating an increased potential for cross-track errors. In the case of Typhoon Songda, elongated ellipses in the RCR and MLR show the influence of variable steering flow and forecasted TC recurvature influenced by mid-latitude flow patterns. This concurs with the ATE statistical analysis for all EPS, which indicated that forecast

positions are behind the actual TC tracks, which is observed in the elongated ellipses after the forecasted point of recurvature in the RCR and MLR the case of Typhoon Songda.

This case study showed that consensus-based forecasts and individual numerical models did not consistently demonstrate value when portraying forecast variability. The use of EPS ellipses conveyed potential for high uncertainty over the cross-track and along-track directions.

## **V. CONCLUSIONS AND RECOMMENDATIONS**

### **A. CONCLUSIONS**

In this thesis, a statistical analysis of uncertainty in TC track forecasts was performed by examining the attributes of ensemble-mean forecasts of TC tracks for these operational EPS. Thus, the goal was to provide added guidance for improvements to the current process of creating operational TC forecasts and visual decision aids.

The ECMWF ensemble has the highest reliability and resolution of the three EPS examined. As discussed, variations in performance of each EPS in individual sub-regions indicate that each EPS has unique attributes. The UKMO has the smallest FTE in the SCSR, which is an example of how the attributes of individual EPS in specific sub-regions could be weighed and exploited to better enhance a deterministic ensemble approach to TC forecasting.

In the case studies, it was evident that consensus-based forecasts and individual numerical models did not demonstrate value portraying forecast variability. The use of EPS ellipses conveyed potential for high uncertainty over the cross-track and along-track directions.

Based on the results of this study, using individual ensemble predictions systems may provide advantages in conveying uncertainty in TC forecasts. The use of the multiple single-model ensembles and anisotropic ellipses designed to represent ensemble spread does enhance spatial and temporal representation of uncertainty, which demonstrates utility with respect to TC track prediction.

### **B. RECOMMENDATIONS**

In the analysis section, the use of a Grand Ensemble (GE) of the three EPS in the ellipse reliability was not performed. Further research in this subject should also include a modification in the MAD analysis using the GPCE circle as the control circle could improve the understanding of resolution.

Further study into significant biases or systematic errors of along-track and across-track errors could reveal sources of uncertainty in speed and space of forecast tracks in specific sub-regions. The use of a GE of all three EPS in the WPAC sub-regions may also help further define the utility of GE ellipses. Additionally, the sub-region boundaries could be improved based on the climatology of storm tracks and storm intensities in the WESTPAC basin, while also using a larger period of data availability.

In addition, the development of forecast uncertainty visual aids using the individual EPS performance in the WPAC, and sub-regions, may enhance forecasts and visual decision aids produced by JTWC.

Lastly, it would be beneficial to conduct decision analysis based on the cost of sortie versus non-sortie and damage vs. non-damage; using contoured product of graded probabilities.



## LIST OF REFERENCES

- AGBOM, cited 2012: Australian Government Bureau of Meteorology (AGBOM) Tropical Cyclone Page [Available online at <http://www.bom.gov.au/cyclone/cxmlinfo/index.shtml>] accessed 28 July 2012.
- Bowler N. E, A. Arribas, K. R. Mylne, K. B. Robertson and S. E. Beare, 2008: The MOGREPS short-range ensemble prediction system. Q. J. Roy. *Meteorol. Soc.*, **134**, 706.
- Buckingham, C., T. Marchok, I. Ginis, L. Rothstein, and D. Rowe, 2010: Short- and medium-range prediction of tropical and transitioning cyclone tracks within the NCEP global ensemble forecasting system. *Weather and Forecasting*, **25**, 1741–1742.
- Carr, L. E. and R. L. Elsberry R. L., 2000: Consensus of dynamical tropical cyclone track forecasts-error versus spread. *Mon. Wea. Rev.* **128**, 4131–4138.
- DeMaria, M., J. A. Knaff, R. Knabb, C. Lauer, C. R. Sampson, and R. T. DeMaria, 2009a: A new method for estimating tropical cyclone wind speed probabilities. *Weather and Forecasting*, **24**, 1574–1575
- ECMWF, cited 2012: Different perturbation techniques [Available online at [http://www.ecmwf.int/products/forecasts/guide/Different\\_perturbation\\_techniques.html](http://www.ecmwf.int/products/forecasts/guide/Different_perturbation_techniques.html)] accessed 20 July 2012.
- Goerss, J S., 2000: Tropical cyclone track forecasting using an ensemble of dynamical models. *Mon. Wea. Rev.*, **128**, 1187–1193
- Goerss, J S., 2007: Prediction of consensus tropical cyclone track forecast error. *Mon. Wea. Rev.*, **135**, 1985–1990
- Hansen, J. A., J. S. Goerss, C. Sampson, 2010: GPCE-AX: An Anisotropic Extension to the Goerss Consensus Error. *Wea. Forecasting*, **26**, 416-422.
- JTWC, cited 2012a: Annual Tropical Cyclone Reports [Available online at <http://www.usno.navy.mil/JTWC/annual-tropical-cyclone-reports>] accessed 4 April 2012.
- JTWC, cited 2012b: JTWC Warning Graphic Legend [Available online at <http://www.usno.navy.mil/JTWC/annual-tropical-cyclone-reports>] accessed 4 April 2012.
- JTWC, cited 2012c: JTWC Collaboration Site [Available online at <https://pzal.nmci.navy.mil/cgi-bin/collab.cgi>] accessed 4 April 2012.

- Lorenz, E. N., 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.*, **42**, 433–471.
- Lorenz, E. N., 1965: A study of the predictability of a 28-variable atmospheric model. *Tellus*, **17**, 321–333.
- Majumbar, S. J. and P. M. Finocchio, 2009: On the ability of global ensemble prediction systems to predict tropical cyclone track probabilities *Wea. Forecasting*, **25**, 659–680.
- Nixon, Christopher, 2012: Evaluating tropical cyclone forecast track uncertainty using a grand ensemble of ensemble prediction systems. M. S. Thesis, Naval Postgraduate School, 60 pp. [Available from <http://edocs.nps.edu/>]
- Pearman, D. W., 2011: Evaluating tropical cyclone forecast track uncertainty using a grand ensemble of ensemble prediction systems. M. S. Thesis, Naval Postgraduate School, 63 pp. [Available from <http://edocs.nps.edu/>]
- Yamaguchi, M. and S. J. Majumdar, 2010: Using TIGGE Data to Diagnose Initial Perturbations and Their Growth for Tropical Cyclone Ensemble Forecasts. *Mon. Wea. Rev.*, **138**, 3634–3655.

## INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center  
Ft. Belvoir, Virginia
2. Dudley Knox Library  
Naval Postgraduate School  
Monterey, California
3. Director  
Joint Typhoon Warning Center  
Pearl Harbor, HI
4. Superintendent  
Naval Research Laboratory  
Monterey, CA
5. Professor Patrick Harr  
Naval Postgraduate School  
Monterey, California
6. Professor Joshua Hacker  
Naval Postgraduate School  
Monterey, California
7. Lieutenant Commander David Marino, United States Navy  
Naval Oceanographic Anti-Submarine Warfare Command  
Naval Base Yokosuka, Japan